Rotman

THE THEORY OF GROUPS: An Introduction Second Edition

ALLYN AND BACON

THE THEORY OF GROUPS

An Introduction

Second Edition

JOSEPH J. ROTMAN

The Theory of Groups

An Introduction

לזכרון נצח אבי מורי הנכבד

אליהו בן יוסף יונה הלוי

Second Edition

THE THEORY OF GROUPS

An Introduction

JOSEPH J. ROTMAN

University of Illinois
Urbana

51120

Allyn and Bacon, Inc. Boston

This book is part of the

ALLYN AND BACON SERIES IN ADVANCED MATHEMATICS

Consulting Editor: IRVING KAPLANSKY

Other books in the series:

James Dugundji, *Topology*
Horst Herrlich and George E. Strecker, *Category Theory*
Irving Kaplansky, *Commutative Rings*
Irving Kaplansky, *Linear Algebra and Geometry: A Second Course*
Irving Kaplansky, *Set Theory and Metric Spaces*
Ralph Kopperman, *Model Theory and Its Applications*

© Copyright 1973 by Allyn and Bacon, Inc.
470 Atlantic Avenue, Boston.

All rights reserved.

No part of the material
protected by this copyright notice
may be reproduced or utilized in any form
or by any means, electronic or mechanical,
including photocopying, recording, or by any
informational storage and retrieval system
without written permission
from the copyright owner.

Library of Congress Catalog Card Number: 72-91980

Printed in the United States of America.

Contents

Preface *vii*

1. GROUPS AND HOMOMORPHISMS **1**

2. THE ISOMORPHISM THEOREMS **11**
Subgroups *11*
Normal Subgroups and Quotient Groups *18*
The Isomorphism Theorems *21*
The Correspondence Theorem *24*

3. PERMUTATION GROUPS **27**
Permutations *27*
Cycles *28*
Conjugates *34*
The Simplicity of A_n *38*
Some Representation Theorems *42*
Counting Orbits *44*

4. FINITE DIRECT PRODUCTS **49**
Direct Products *49*
The Basis Theorem *52*
The Fundamental Theorem of Finite Abelian Groups *57*
Modules and Matrices *60*
The Remak-Krull-Schmidt Theorem *75*

5. THE SYLOW THEOREMS **83**
p-Groups *83*
The Sylow Theorems *85*
Some Applications of the Sylow Theorems *89*

6. NORMAL SERIES **96**
Some Galois Theory *96*
The Jordan-Hölder Theorem *103*
Solvable Groups *108*
A Theorem of P. Hall *112*
Central Series and Nilpotent Groups *116*

7. EXTENSIONS **127**
The Extension Problem *127*
Automorphism Groups *129*
Semidirect Products *135*
Factor Sets *141*
The Schur-Zassenhaus Lemma *148*

8. SOME SIMPLE GROUPS **152**
Finite Fields 152
The General Linear Group 156
PSL(2, K) 160
PSL(m, K) 165

9. INFINITE ABELIAN GROUPS **174**
The First Reduction: Torsion and Torsion-Free 174
The Second Reduction: Divisible and Reduced 179
Free Abelian Groups 187
Finitely Generated Abelian Groups 192
Torsion Groups 194
Torsion-Free Groups 200
Grothendieck Groups 204

10. HOMOLOGICAL ALGEBRA **211**
The Hom Functors 211
Definition of Ext 217
Pull-Backs and Push-Outs 219
The Ext Functors 228

11. FREE GROUPS AND FREE PRODUCTS **236**
Generators and Relations 236
Edgepath Groups 243
Covering Complexes and the Nielsen-Schreier Theorem 251
Free Products and the Kuroš Theorem 262
Free Products with Amalgamated Subgroup 269

12. THE WORD PROBLEM **277**
Statement of the Problem 277
Turing Machines and Post's Theorem 279
Britton's Lemma 289
The Novikov-Boone Theorem 298
An Imbedding Theorem of G. Higman 308

APPENDICES **322**
I *Some Major Algebraic Systems* 322
II *Equivalence Relations and Equivalence Classes* 324
III *Functions* 326
IV *Zorn's Lemma* 328
V *Principal Ideal Domains* 330

Bibliography 333
Notation 335
Index 337

Preface

Quand j'ai voulu me restreindre, je suis tombé dans l'obscurité; j'ai préféré passer pour un peu bavard.

H. POINCARÉ, *Analysis situs*, Journal de l'Ecole Polytechnique, 1895, pp. 1–121.

This text is designed for the student who wants to learn the basic ideas and techniques of the theory of groups. We assume that the reader knows the rudiments of modern algebra, by which we mean that matrices and finite-dimensional vector spaces are friends, while groups, rings, fields, and homomorphisms are only acquaintances. A familiarity with elementary set theory is also assumed, but some appendices are at the back of the book so that the reader may see whether my notation is the same as his.

No claims are made that this is an encyclopedia of elementary group theory. In fact, many important topics, e.g., representation theory, transfer, and topological groups, are not treated. The omission of representation theory, an indispensable tool of any group theorist, may seem to be a particularly serious omission, but I feel that a considerable amount of ring theory must be developed to present this subject properly, and limitations of space prevent such a development here. Moreover, an excellent exposition of representation theory already exists in *Representation Theory of Finite Groups and Associative Algebras*, by C. Curtis and I. Reiner.

The table of contents gives an outline of what has not been omitted. Here are some comments on unusual aspects of each chapter. 1. *Groups and Homomorphisms;* 2. *The Isomorphism Theorems.* Since I believe that categories are useful, I have tried to emphasize homomorphisms as well as groups whenever reasonable. 3. *Permutation Groups.* After proving the standard theorems (unique factorization into disjoint cycles, simplicity of A_n, Cayley's theorem), we give an elementary result (probably due to Burnside) which Polya has applied in solving some difficult counting problems. 4. *Finite Direct Products.* The proofs of the basis theorem and the fundamental theorem for finite abelian groups have been written so that the proofs of their analogs for modules over principal ideal domains are easily seen. These analogs are applied to prove existence and uniqueness of canonical forms for matrices. The chapter ends with the Remak-Krull-Schmidt theorem (I am indebted to M. Barr for suggesting an improvement in my original account of this theorem, as well as for suggesting Exercises 3.41,

3.45, and 4.3). 5. *The Sylow Theorems.* Sylow subgroups are defined as maximal p-subgroups, so that existence and imbedding are immediate consequences of the definition. 6. *Normal Series.* A brief introduction to Galois theory is given, so that the reader may see the genesis of our subject. Although spelling reformers may have good intentions, they should realize that they too often introduce a cure for a disease that is not really pernicious. Accordingly, I no longer write "subnormal series" when the rest of the world writes "normal series". The chapter contains the usual first theorems on solvable and nilpotent groups, including P. Hall's theorem that finite solvable groups have p-complements, and ends with the classification of finite p-groups having a unique subgroup of order p. 7. *Extensions.* Several automorphism groups are computed, semidirect products with nonabelian kernels are discussed, and wreath product is introduced and used to compute sylow subgroups of symmetric groups. Only when nonsplit extensions are discussed do we insist that kernels be abelian, in which case we use Mac Lane's "logarithmic" notation: if $G/K \cong Q$, then Q is written multiplicatively while G and its subgroup K are written additively. The theory is sufficiently developed to prove the Schur-Zassenhaus lemma. 8. *Some Simple Groups.* The groups *PSL* are proved, in the main, to be simple by geometric methods (after E. Artin; our treatment follows notes of F. D. Veldkamp). However, a matrix proof is given for the 2×2 case, for I think this is easier. Two nonisomorphic simple groups of the same order are exhibited in the chapter's final section. 9. *Infinite Abelian Groups.* Since injectivity is more intuitive than projectivity, divisible groups are treated before free abelian groups. We give two proofs of the subgroup theorem, the first working only when the group is of finite rank, for any transfinite proof (which is needed, of course, in the general case) disguises the simple underlying idea. A new section on Grothendieck groups should serve as a gentle introduction to algebraic K-theory; more important, it formalizes the heuristic treatment of extensions as "products" by actually constructing a group whose elements are (classes of) groups and whose binary operation is extension. This leads to elegant descriptions of the Remak-Krull-Schmidt and Jordan-Hölder theorems. 10. *Homological Algebra.* The basic properties of Ext_Z are proved by considering both factor sets and classes of short exact sequences. Although the reader cannot appreciate the power of homological methods from this little chapter, my intention is to show that this subject does arise quite naturally out of group theory and, even in the beginning, does say something of interest. 11. *Free Groups and Free Products.* The notion of generators and relations motivates the definition of a free groups as a solution to a universal mapping problem. The Nielsen-Schreier subgroup theorem is proved by using edgepath groups of abstract simplicial complexes (so this proof is thus an algebraic adaptation of the "topological proof" of Baer and Levi). This takes a little longer than a direct assault using cancellation arguments, but it has the advantage of being better understood than the standard proofs. An added bonus is that the Kuroš subgroup theorem may be proved with very little additional work. (I owe much to the account of the topological theory found in the beautiful book of Massey.) The proof we give of Theorem 11.37 is due to P. Schupp. The treatment of free product with

amalgamation is motivated by its analog with union; indeed, it is even possible to mention van Kampen's theorem. We treat only the special case of free product of two groups with an amalgamation rather than the general case of many factors with amalgamation, for it is easier to describe and is the most useful such construction. 12. *The Word Problem.* This chapter is the most unusual section in the book. Aside from the customary license given to an author for his closing chapter, there are several reasons for discussing the word problem: It is an interesting problem related to groups and the solution of W. W. Boone and J. L. Britton has a group-theoretical flavor. I have taken the advice of many in rewriting the final chapter. A more complete proof of the Post-Markov theorem is given (the only gap being an appeal to Church's thesis in the midst of giving a nonrecursive r.e. set). A modification, due to C. F. Miller, III, of Britton's proof of the Novikov-Boone theorem shortens my original presentation (C. Jockusch and T. McLaughlin also made helpful comments on this material). We end with S. Aanderaa's new proof of G. Higman's characterization of the finitely generated subgroups of finitely presented groups (D. Collins and C. F. Miller, III, kindly provided me with excellent notes explicating Aanderaa's proof). It is a pleasure to give special thanks to K. I. Appel, W. W. Boone, J. L. Britton, and P. Schupp for their patience and enthusiastic help with the last chapter.

I am fortunate in having attended lectures on group theory given by I. Kaplansky, S. Mac Lane, and M. Suzuki. Their influence is evident throughout in many elegant ideas and proofs. C. R. B. Wright read my original manuscript and made many valuable comments. I wish to thank also my colleagues at the University of Illinois, especially G. Brown, J. Eagon, and H. Paley, for their helpful suggestions with the first edition. The National Science Foundation supported a part of this work and I gratefully acknowledge their aid.

Finally, I thank Mrs. Divona Keel for a superb job of typing my manuscript.

Urbana, 1971 JOSEPH J. ROTMAN

To the Reader

Exercises in a text generally have two functions: to reinforce the reader's grasp of the material; to provide puzzles whose solution gives a certain pleasure. In this text, the exercises have a third function: to enable the reader to discover important facts, examples, and counterexamples. The serious reader should attempt the exercises as they arise (many are not difficult), for subsequent proofs often depend on them. The casual reader should regard the exercises as part of the text proper. Those exercises preceded by * will be cited elsewhere in the book; those preceded by ** will be cited in a proof.

CHAPTER 1

Groups and Homomorphisms

Our study of groups begins with the consideration of a rather general setting.

Definition If G is a nonempty set, a **binary operation** on G is a function $\mu: G \times G \to G$.

A binary operation assigns to each ordered pair a, b of elements of G a third element $\mu(a, b)$ of G. In practice, μ is thought of as a multiplication of elements of G, and instead of $\mu(a, b)$, more suggestive notations used are ab, $a + b$, $a \circ b$, and $a * b$. In this chapter, we shall use the latter notation.

Two remarks must be made here. First, it is quite possible that $a * b$ and $b * a$ are distinct elements of G. Second, the law of substitution—if $a = a'$ and $b = b'$, then $a * b = a' * b'$—is just the statement that μ is a well-defined (i.e., single-valued) function.

One cannot develop a theory in this rarefied atmosphere; conditions on the binary operation are needed in order to obtain interesting results. If we are given elements a_1, a_2, a_3 of G (not necessarily distinct), the notation $a_1 * a_2 * a_3$ is ambiguous. Since one can $*$ only two elements of G at a time, there is a choice: form $a_1 * a_2$ first, and then $*$ this new element of G with a_3 to get $(a_1 * a_2) * a_3$; or form $a_1 * (a_2 * a_3)$. In general, these two elements are not the same. For example, if G is the set of all integers (positive, negative, and zero) and if the binary operation

is subtraction, i.e., $a * b = a - b$, then almost any choice of integers a, b, c yields an example in which $(a - b) - c \neq a - (b - c)$.

Definition A binary operation $*$ on G is **associative** in case

$$(a * b) * c = a * (b * c)$$

for every three elements a, b, c of G.

Associativity allows one to multiply every ordered triple of elements of G unambiguously; parentheses are unnecessary, and there is no confusion in writing $a * b * c$. If we are confronted by four elements of G or, more generally, by a finite number of elements of G, must we postulate more intricate associativity axioms in order to do without parentheses?

Theorem 1.1 *Let $*$ be an associative binary operation on G. Any two ways of multiplying the elements $a_1, a_2, \cdots, a_n$ of G in this order yield the same element of G.*

Proof We prove the theorem by induction on n, where $n \geq 3$. If $n = 3$, the theorem is true, for we are assuming that $*$ is associative.

Let $n > 3$. How does one multiply n elements? Since a binary operation allows multiplication of only two elements at a time, one first chooses a pair of adjacent a's and multiplies them; then, either he multiplies this new element by an a adjacent to it or he chooses two other adjacent a's and multiplies them. With each choice, the total number of factors decreases by one, so that eventually only two factors are left. Suppose now that

$$X = (a_1 * \cdots * a_i) * (a_{i+1} * \cdots * a_n)$$

and

$$Y = (a_1 * \cdots * a_j) * (a_{j+1} * \cdots * a_n)$$

are elements of G obtained by two people multiplying the a's together, each having made his own choices. The parentheses indicate the final multiplication each has just performed. For notational convenience, we assume that $i \leq j$. Since each of the final factors in X and in Y contains less than n of the a's, the inductive hypothesis allows us to rearrange parentheses in them. Therefore, we may assume $i < j$ (or we are done) and write

$$X = (a_1 * \cdots * a_i) * ([a_{i+1} * \cdots * a_j] * [a_{j+1} * \cdots * a_n])$$

and

$$Y = ([a_1 * \cdots * a_i] * [a_{i+1} * \cdots * a_j]) * (a_{j+1} * \cdots * a_n).$$

If we denote $a_1 * \cdots * a_i$ by A, $a_{i+1} * \cdots * a_j$ by B, and $a_{j+1} * \cdots * a_n$ by C, then $X = A * (B * C)$ and $Y = (A * B) * C$. Since A, B, and C are unambiguously defined, by induction, and since $*$ is associative, we have $X = Y$, as we wished. ∎

As a result of this theorem, the notation $a_1 * a_2 * \cdots * a_n$ is unambiguous when $*$ is an associative multiplication.

Definition A **semigroup** is a set G with an associative binary operation.

A semigroup is a pair, but we shall usually say that G is a semigroup and tacitly assume that $*$ is known. The reader must realize, however, that there are many possible operations on a set G that make it a semigroup. For example, the set of positive integers is a semigroup under either of the binary operations of ordinary addition or ordinary multiplication.

Definition Let G be a semigroup and let $a \in G$. Let $a^1 = a$; if $n > 1$, then $a^n = a^{n-1} * a$.

Corollary 1.2 *Let G be a semigroup and let $a \in G$. If m and n are positive integers, $a^m * a^n = a^{m+n}$*

Proof Both elements arise from a product of $m + n$ factors equal to a; they differ only in the arrangement of parentheses. ∎

The most important semigroups are groups.

Definition A **group** is a semigroup G containing an element e such that:

(i) $e * a = a$ for all $a \in G$;
(ii) for every $a \in G$, there is an element b in G with

$$b * a = e.$$

Lemma 1.3 *If G is a group and $a \in G$, then $a * a = a$ implies $a = e$.*

Proof There is an element $b \in G$ with $b * a = e$. Now

$$b * (a * a) = b * a = e.$$

On the other hand,

$$(b * a) * a = e * a = a.$$

Therefore, $a = e$. ∎

Theorem 1.4 *In a group G, $a * e = a$ for all $a \in G$, and if $b * a = e$, then $a * b = e$. Furthermore, there is only one element e satisfying* (i), *and given any $a \in G$, there is only one $b \in G$ satisfying $b * a = e$.*

Proof We first show that in this particular case, the order in which we multiply makes no difference.

Suppose $b * a = e$. Then $(a * b) * (a * b) = a * (b * a) * b = a * (e * b) = a * b$. By Lemma 1.3, $a * b = e$.

If $a \in G$, then $e * a = a$. Choose b with $b * a = e$. Then $a * e = a * (b * a) = (a * b) * a = e * a = a$, by our calculation above. Therefore $a * e = a$.

We now prove the uniqueness assertions. Suppose $e_0 * a = a$ for all $a \in G$; in particular, $e_0 * e_0 = e_0$. By the lemma, $e_0 = e$.

Finally, suppose $b * a = e$ and $c * a = e$. Then $a * b = e$ and $c = c * e = c * (a * b) = (c * a) * b = e * b = b$. ∎

As a result of the uniqueness assertions of the theorem, we may now give names to e and b. We call e the **identity** of G, and if $b * a = e$, we call b the **inverse** of a and denote it a^{-1}.

Corollary 1.5 If G is a group and $a \in G$, then

$$(a^{-1})^{-1} = a.$$

Proof $(a^{-1})^{-1}$ is that element $b \in G$ such that $b * a^{-1} = e$. But we have just seen that $a * a^{-1} = e$, so that the uniqueness of the inverse implies that $b = a$. ∎

Definition Let G be a group and $a \in G$. We define the **powers** of a as follows: If n is a positive integer, then a^n has been defined above; define $a^0 = e$; finally, define a^{-n} as $(a^{-1})^n$.

EXERCISES

1.1. If G is a group and $a_1, \cdots, a_n \in G$, then

$$(a_1 * \cdots * a_n)^{-1} = a_n^{-1} * \cdots * a_1^{-1}.$$

1.2. Let G be a group, $a \in G$, and m, n possibly negative integers. The following formulas hold:

$$a^m * a^n = a^{m+n};$$

$$(a^m)^n = a^{mn}.$$

1.3. **(Cancellation laws)** In a group G, each of the equations $a * b = a * c$ and $b * a = c * a$ implies $b = c$.

1.4. A semigroup G is a group if and only if the equations $a * x = b$ and $y * a = b$ have solutions in G for every a and b in G. Moreover, when G is a group, these solutions are unique.

****1.5.** Let G be a group and let $a \in G$. Define a function $T_a: G \to G$ by $T_a(x) = a * x$ (T_a is called **left translation** by a). Prove that T_a is a one-to-one correspondence.

1.6. Let Z denote the additive group of integers (we shall adhere to this notation from now on). What is the identity of Z? If $a \in Z$, what is its inverse? State Corollary 1.5 for Z.

1.7. For each fixed $n > 0$, prove that Z_n, the set of integers modulo n, is a group if one defines $\bar{a} + \bar{b} = \overline{a + b}$ ($\bar{a}$ is the congruence class of a modulo n, i.e.,

$$\bar{a} = \{b \in Z: b \equiv a(\text{mod } n)\} = \{a + kn: k \in Z\}).$$

Conclude that, for each $n > 0$, there exists a group having exactly n elements.

1.8. Let G denote the multiplicative group of positive rationals. What is the identity of G? If $a \in G$, what is its inverse?

1.9. Let n be a positive integer, and let G be the multiplicative group of all nth roots of unity, i.e., all complex numbers of the form $e^{(2\pi ik)/n}$, where $k \in Z$. What is the identity of G? If $a \in G$, what is its inverse? How many elements does G have?

1.10. Let M be the multiplicative group of all 2×2 nonsingular matrices with rational entries. What is the identity of M? If $a \in M$, what is its inverse? Exhibit elements a and b in M such that $a * b \neq b * a$.

***1.11.** Prove that the following four matrices form a multiplicative group:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

***1.12.** Let X be a nonempty set and let G be the set of all one-to-one correspondences of X onto itself. If f and g are in G, define $f * g$ to be the composite of f and g:

$$(f * g)(x) = f(g(x))$$

for every $x \in X$. Prove that G is a group. (If $X = \{1, 2, \cdots, n\}$, then G is denoted S_n.)

When does one "know" a group G? One answer (but not the only answer, as we shall see) is that all the elements of G are known and all possible products can be computed. This answer can be made to look more sophisticated in the special case when G is finite. Let $a_1, a_2, \cdots, a_n$ be a list with no repetitions of all the elements in G. A **multiplication table** of G is the $n \times n$ matrix whose ij entry is $a_i * a_j$:

G	a_1	a_2	$\cdots$	a_n
a_1	$a_1 * a_1$	$a_1 * a_2$	$\cdots$	$a_1 * a_n$
a_2	$a_2 * a_1$	$a_2 * a_2$	$\cdots$	$a_2 * a_n$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
a_n	$a_n * a_1$	$a_n * a_2$	$\cdots$	$a_n * a_n$

Therefore, one "knows" a finite group if he can write a multiplication table for it. (We say "a" multiplication table rather than "the" multiplication table, since the matrix depends on the particular listing of the elements of G.) One also speaks of a multiplication table of an infinite group, but in this case, of course, the matrix is infinite.

Let us now consider two almost trivial examples of groups. Let G be the multiplicative group with elements $\{1, -1\}$; let $H = Z_2$, the additive group of integers modulo 2. Compare multiplication tables of these two groups:

G	1	-1
1	1	-1
-1	-1	1

H	$\bar{0}$	$\bar{1}$
$\bar{0}$	$\bar{0}$	$\bar{1}$
$\bar{1}$	$\bar{1}$	$\bar{0}$

It is quite clear that G and H are distinct groups; it is equally clear that there is no significant difference between them. We formalize this idea.

Definition Let $(G, *)$ and $(H, \circ)$ be groups. A **homomorphism** $f: G \to H$ is a function for which

$$f(a * b) = f(a) \circ f(b) \qquad \text{for all } a, b \in G.$$

An **isomorphism** is a homomorphism that is also a one-to-one correspondence.[1] Two groups G and H are **isomorphic**, denoted $G \cong H$, if there is an isomorphism $f: G \to H$.

[1] Many authors use "isomorphism" to mean "homomorphism and one-to-one"; here, isomorphisms are always onto as well.

The groups G and H, whose multiplication tables are given above, are isomorphic: define $f: G \to H$ by $f(1) = \bar{0}$ and $f(-1) = \bar{1}$. It is important to realize that a given group has many multiplication tables, one for each way of listing the elements of the group. In particular, a list $x_1, x_2, \cdots$ of the elements of a group G determines a multiplication table for G. If $f: G \to H$ is a one-to-one correspondence, then $f(x_1), f(x_2), \cdots$ determines a multiplication table for H. To say that f is an isomorphism is to say that, upon superimposing the first multiplication table (for G) on the second table (for H), the tables "match". In this sense, one abuses language and says that isomorphic groups have the same multiplication table.

EXERCISES

1.13. Write a multiplication table for S_3 (see Exercise 1.12).

1.14. Isomorphic groups have the same number of elements. Prove that the converse is false by showing that Z_4 is not isomorphic to the group of four matrices in Exercise 1.11.

1.15. For each positive integer n, prove there are only finitely many distinct groups having exactly n elements (we regard isomorphic groups as being the same).

Two basic problems occurring in mathematics are: (1) the classification of all systems of a given kind, e.g., all groups, all vector spaces, all topological spaces; and (2) the classification of all the transformations of one system into another. By a classification of systems, one usually means a scheme that distinguishes essentially different systems, or to say it another way, a scheme that tells when two systems are essentially the same. A classification of transformations is more subtle, and we needn't discuss it further now. As an illustration, consider the collection of all finite-dimensional vector spaces over a field F. In this case, the first problem is answered by the theorem that two such spaces are isomorphic if and only if they have the same dimension. Even the second problem has been answered. The transformations of vector spaces are, of course, the linear transformations, which give rise to similarity classes of matrices, and these are classified by the canonical forms. The same problems arise in group theory: (1) when are two groups isomorphic; (2) how may one describe the homomorphisms from one group to another? In contrast to our illustration, both problems are exceedingly difficult (if not impossible) and are only partially solved.

The reader is advised that homomorphisms are powerful tools in the study of groups, and they should be used as much as possible.

EXERCISES

1.16. If $f: G \to H$ and $g: H \to K$ are homomorphisms, then so is $g \circ f: G \to K$.

1.17. If $\mathscr{A}$ is a collection of groups, then the relation $G \cong H$ is an equivalence relation on $\mathscr{A}$.

***1.18.** Let $f: G \to H$ be a homomorphism, and let $a \in G$. For every $n \in Z, f(a^n) = f(a)^n$. Note the special cases $n = 0$ and $n = -1$.

****1.19.** Let G be a group and let X be a set having the same number of elements as G. If $f: G \to X$ is a one-to-one correspondence, there is a unique binary operation that can be defined on X so that X is a group and f is an isomorphism.

1.20. Let G be the multiplicative group of positive reals and let H be the additive group of all reals. Show that $\log: G \to H$ is an isomorphism. (HINT: Exhibit a function inverse to log.)

***1.21.** Let G be the additive group of reals and let $\mathbf{T}$ be the **circle group**, i.e., the multiplicative group of all complex numbers of absolute value 1. For a fixed real number y, prove that $f: G \to \mathbf{T}$ defined by $f(x) = e^{iyx}$ is a homomorphism. Which real numbers x are such that $f(x) = 1$? (It can be proved that these f are the only *continuous* homomorphisms from G to $\mathbf{T}$.)

1.22. If G is the multiplicative group of all $n \times n$ nonsingular matrices with entries in a field F, and if F^* is the multiplicative group of nonzero elements of F, then determinant is a homomorphism from G to F^*.

1.23. Let a be a fixed element of a group G. Prove that $f: G \to G$ defined by

$$f(x) = a * x * a^{-1}$$

is an isomorphism (f is called **conjugation** by a). We often write $a * x * a^{-1} = x^a$.

1.24. Let G be the multiplicative group of all nth roots of unity; prove that $G \cong Z_n$.

1.25. Let G be the additive group of integers modulo 6. Describe all the homomorphisms from G to itself. Which of these are isomorphisms?

*1.26. Let G be the additive group of $Z[x]$, all polynomials in x having coefficients in Z, and let H be the multiplicative group of positive rationals. Prove that $G \cong H$. (HINT: Use the fundamental theorem of arithmetic to construct an isomorphism.)

Having solved Exercise 1.26, the reader may wish to reconsider his answer to the question: When does one "know" a group? Surely the reader could construct multiplication tables for the two groups G and H, but it was not obvious at the outset that these groups are essentially the same. As an alternative answer to the question of "knowing", we suggest that one knows a group G when he can determine, given any other group H, whether G and H are isomorphic.

Definition Two elements a and b in a semigroup G **commute** in case $a * b = b * a$. A semigroup is **abelian**[2] (or commutative) in case every two elements in G commute.

EXERCISES

1.27. Let $a_1, a_2, \cdots, a_n$ be elements of an abelian semigroup G. If $b_1, b_2, \cdots, b_n$ is a rearrangement of the a_i, then

$$a_1 * a_2 * \cdots * a_n = b_1 * b_2 * \cdots * b_n.$$

1.28. Let a and b lie in a semigroup G. If a and b commute, then $(a * b)^n = a^n * b^n$ for every integer $n > 0$. If, further, G is a group, the equation holds for all $n \in Z$.

1.29. Let G be a group in which the square of every element is the identity. Prove that G is abelian.

1.30. A group G is abelian if and only if the function $f: G \to G$ defined by $f(x) = x^{-1}$ is a homomorphism.

Definition A semigroup G is a **cancellation semigroup** in case, for every $a, b, c \in G$, each of the equations $a * b = a * c$ and $b * a = c * a$ implies $b = c$.

EXERCISES

1.31. Prove that every finite cancellation semigroup is a group.

1.32. Prove that every abelian cancellation semigroup G can be imbedded in a group. (HINT: As in the construction of the rationals

[2] In honor of N. H. Abel (1802–1829).

from the integers, consider all ordered pairs $(a, b) \in G \times G$ under the equivalence relation:

$$(a, b) \equiv (a', b') \qquad \text{if } a * b' = a' * b.)$$

Mal'cev has exhibited a cancellation semigroup that cannot be imbedded in a group (see P. M. Cohn, *Universal Algebra*, p. 269).

1.33. **(Kaplansky)** An element x in a ring R has a *left quasi-inverse* in case there is an element $y \in R$ with $x + y - yx = 0$. Prove that a ring R in which every element but one has a left quasi-inverse is a division ring. (HINT: Let R' be the set R with the exceptional element deleted; consider the binary operation on R' defined by $x \circ y = x + y - yx$.)

1.34. Let G be a finite abelian group containing no elements $a \neq e$ with $a^2 = e$. Evaluate $a_1 * \cdots * a_n$, where $a_1, \cdots, a_n$ is a list with no repetitions of all the elements of G.

1.35. Prove Wilson's theorem: If p is a prime, then

$$(p - 1)! \equiv -1 \pmod{p}.$$

(HINT: The nonzero elements of the integers modulo p form a group under multiplication.)

***1.36.** Let G be a finite group and $f: G \to G$ an isomorphism. If f has no fixed points (i.e., $f(x) = x$ implies $x = e$) and if $f \circ f$ is the identity function on G, then G is abelian. (HINT: Prove that every element in G has the form $x^{-1} * f(x)$.)

CHAPTER 2

The Isomorphism Theorems

SUBGROUPS

From now on, we shall usually write ab instead of $a * b$, and 1 will denote the identity element instead of e.

Definition A nonempty subset S of a group G is a **subgroup** of G in case S is a group under the binary operation of G.

Let G be the additive group of reals and S the multiplicative group of positive rationals. S is a group contained in G, but S is not a subgroup, since the binary operations are not the same. Indeed, a group S contained in G is a subgroup if and only if the inclusion $i\colon S \to G$ is a homomorphism. It follows from Exercise 1.18 that the identity elements of S and of G are the same.

Theorem 2.1 *If S is a subset of a group G, then S is a subgroup of G if and only if:*

(i) $1 \in S$;
(ii) $a \in S$ implies $a^{-1} \in S$;
(iii) $a, b \in S$ implies $ab \in S$.

Proof The verifications of the axioms are easy. ∎

Theorem 2.2 *If S is a subset of a group G, then S is a subgroup of G if and only if S is nonempty, and whenever $a, b \in S$, then $ab^{-1} \in S$.*

Proof We use Theorem 2.1.

(i) Since S is nonempty, it contains an element a. Hence, $aa^{-1} = 1 \in S$.
(ii) If $b \in S$, then $1b^{-1} = b^{-1} \in S$.
(iii) If $a, b \in S$, then $a(b^{-1})^{-1} = ab \in S$. ∎

Example 1 If G is a group, G itself and $\{1\}$ are subgroups of G; any other subgroup of G is called **proper**.

Example 2 If $a \in G$, let $\langle a \rangle$ denote the set of all powers of a. $\langle a \rangle$ is a subgroup, the **cyclic subgroup generated by** a; a is called a **generator** of $\langle a \rangle$. A cyclic subgroup may have several different generators, e.g., $\langle a \rangle = \langle a^{-1} \rangle$.

Definition A group G is **cyclic** if $G = \langle a \rangle$ for some $a \in G$.

Example 3 Let G and H be groups and let $f: G \to H$ be a homomorphism. Define

$$\textbf{kernel } f = \{x \in G : f(x) = 1\}.$$

We use Theorem 2.1 to show that kernel f is a subgroup of G. Kernel f is nonempty because it contains 1; if $f(x) = 1$, then $f(x^{-1}) = f(x)^{-1} = 1$, so that $x \in$ kernel f implies $x^{-1} \in$ kernel f; if $f(x) = 1 = f(y)$, then $f(xy) = f(x)f(y) = 1 \cdot 1 = 1$, so that $x, y \in$ kernel f implies $xy \in$ kernel f.

Example 4 If $f: G \to H$ is a homomorphism, define

$$\textbf{image } f = \{y \in H : y = f(x) \quad \text{for some } x \in G\}.$$

The reader may check that image f is a subgroup of H.

Notation We will often write $\ker f$ and $\operatorname{im} f$ instead of kernel f and image f.

We now consider several ways of manufacturing subgroups.

Theorem 2.3 *The intersection of any family of subgroups of G is again a subgroup of G.*

Proof Let $\{S_i : i \in I\}$ be a family of subgroups of G and let $S = \bigcap S_i$. Since $1 \in S_i$ for every i, $1 \in S$, and so S is nonempty. Suppose $a, b \in S$.

These elements got into S by being in every S_i. Since each S_i is a subgroup, $ab^{-1} \in S_i$ for every i, and so $ab^{-1} \in S$. ▮

Theorem 2.4 *If X is a subset of a group G, there is a smallest subgroup H of G that contains X, i.e., if S is any other subgroup containing X, then $S \supset H$.*

Proof There do exist subgroups of G containing X, e.g., G itself. Let H be the intersection of all the subgroups of G that contain X. By the preceding theorem, H is a subgroup and it clearly contains X. The fact that H is the smallest such subgroup follows from the observation that the intersection is contained in each of the sets being intersected. ▮

Definition If X is a subset of a group G, the smallest subgroup of G containing X is denoted $\langle X \rangle$ and is called the **subgroup of G generated by X**. We say that X **generates** $\langle X \rangle$.

If X is a finite set, then $X = \{a_1, a_2, \cdots, a_n\}$. For notational convenience, we write $\langle X \rangle = \langle a_1, a_2, \cdots, a_n \rangle$ instead of $\langle \{a_1, a_2, \cdots, a_n\} \rangle$.

EXERCISES

***2.1.** If X is empty, $\langle X \rangle = \{1\}$; if X is nonempty, $\langle X \rangle$ is the set of all finite products of powers of elements in X. In particular, if $X = \{a\}$, then $\langle X \rangle = \langle a \rangle$ is a cyclic subgroup.

2.2. The set-theoretical union of two subgroups is a subgroup if and only if one is contained in the other. Is this true if we replace "two subgroups" by "three subgroups"?

2.3. If H and K are subgroups of G, we denote $\langle H \cup K \rangle$ by $H \vee K$. Prove that $H \vee K$ is the smallest subgroup of G containing H and K.

2.4. Let S be a subgroup of G with $S \neq G$. If $G - S$ is the complement of S in G, then $\langle G - S \rangle = G$.

2.5. The multiplicative group of positive rationals is generated by all rationals of the form $1/p$, where p is prime.

Definition Let S and T be nonempty subsets of a group G.

$$ST = \{st : s \in S \text{ and } t \in T\}.$$

In particular, if we write t for $\{t\}$, then

$$St = \{st : s \in S\}.$$

Definition Let S be a subgroup of G. A **right coset** of S in G is a subset St (a **left coset** is tS). We say that t is a **representative** of St (and also of tS).

EXERCISES

**2.6. The multiplication of nonempty subsets of G defined above is associative.

**2.7. Let S be a subgroup of G. Prove that $St = S$ if and only if $t \in S$.

2.8. If S is a subgroup of G, then $SS = S$.

2.9. If S is a finite nonempty subset of G and $SS = S$, then S is a subgroup of G. Is this true if S is infinite?

**2.10. There is a one-to-one correspondence between S and St; conclude that any two right cosets of S have the same number of elements.

*2.11. **(Product formula)** If X, Y are nonempty subsets of G, then

$$|XY|\,|X \cap Y| = |X|\,|Y|.$$

A right coset St has many representatives, namely, any element st, where $s \in S$. The next lemma gives a criterion for determining whether two right cosets of S are the same when a representative of each is known.

Lemma 2.5 *Let S be a subgroup of G. Then $Sa = Sb$ if and only if $ab^{-1} \in S$ ($aS = bS$ if and only if $a^{-1}b \in S$).*

Proof If $Sa = Sb$, then $a \in Sb$ and $a = sb$ for some $s \in S$. Therefore, $ab^{-1} = s$ and $ab^{-1} \in S$. Conversely, if $ab^{-1} = s \in S$, then $a = sb$, so that $Sa = Ssb = Sb$, by Exercise 2.7. ∎

Theorem 2.6 *If S is a subgroup of G, then any two right cosets of S in G are either identical or disjoint.*

Proof We show that if there is an element x in $Sa \cap Sb$, then $Sa = Sb$. Such an element has the form $x = s_1a = s_2b$, where s_1 and s_2 lie in S. Hence $ab^{-1} = s_1^{-1}s_2 \in S$, so that Lemma 2.5 gives $Sa = Sb$. ∎

Theorem 2.6 may be paraphrased to say that a subgroup S induces a partition of the group G (into right cosets). This being true, there must be an equivalence relation on G lurking somewhere in the background.

EXERCISE

2.12. If S is a subgroup of G, define a relation on G by $a \equiv b$ in case $ab^{-1} \in S$. Prove that this relation is an equivalence relation whose equivalence classes are the right cosets of S. Use this fact to give an alternative proof of Theorem 2.6.

It is interesting to note that each of the three properties of a subgroup listed in Theorem 2.1 yields a corresponding property of the relation $a \equiv b$.

In Exercise 1.13, we asked the reader to give a multiplication table for the group $G = S_3$. The six elements of S_3 may be denoted $\{1, a, b, b^2, ab, ab^2\}$, where $a^2 = 1 = b^3$ and $ba = ab^2$. If $H = \langle a \rangle = \{1, a\}$, then the right cosets of H in S_3 are

$$H = \{1, a\}, \qquad Hb = \{b, ab\}, \qquad Hb^2 = \{b^2, ab^2\}.$$

Note that the left cosets of H in S_3 are

$$H = \{1, a\}, \qquad bH = \{b, ba\} = \{b, ab^2\}, \qquad b^2H = \{b^2, ab\}.$$

Thus, right cosets and left cosets need not coincide (for example, $Hb \neq bH$ in this case).

Theorem 2.7 *If S is a subgroup of G, the number of right cosets of S in G equals the number of left cosets of S in G.*

Proof Let $\mathscr{R}$ denote the set of all right cosets of S in G, and let $\mathscr{L}$ denote the set of all left cosets. We exhibit a one-to-one correspondence $f: \mathscr{R} \to \mathscr{L}$. If $Sa \in \mathscr{R}$, your first guess is to define $f(Sa) = aS$, but this does not work. Your second guess, $f(Sa) = a^{-1}S$, is correct. It must be verified that f is well defined, i.e., if $Sa = Sb$, then $a^{-1}S = b^{-1}S$ (this is why your first guess is wrong). We also leave to the reader the verification that f is a one-to-one correspondence. ∎

Definition If S is a subgroup of G, the **index** of S in G, denoted $[G:S]$, is the number of right cosets of S in G.

Theorem 2.7 tells us that there is no need to define a "right index" and a "left index", for the number of right cosets is the same as the number of left cosets.

Definition The **order** of G, denoted $|G|$, is the number of elements in G.

Theorem 2.8 (Lagrange) *If S is a subgroup of a finite group G, then $[G:S] = |G|/|S|$ (so that $|S|$ divides $|G|$).*

Proof The right cosets of S partition G into $[G:S]$ parts, each of which has precisely $|S|$ elements, by Exercise 2.10. Therefore $|G| = [G:S]|S|$. ∎

Corollary 2.9 *If $|G| = p$, where p is prime, then G is cyclic.*

Proof Since p is prime, Lagrange's theorem says that G can have no proper subgroups. Choose $a \in G$, $a \neq 1$. Then $\langle a \rangle$ is a subgroup of G and $\langle a \rangle \neq \{1\}$, so that $\langle a \rangle = G$. ∎

Definition If $a \in G$, the **order** of a is $|\langle a \rangle|$. (Thus, the order of a is either a positive integer or infinity.)

Corollary 2.10 *If G is a finite group and $a \in G$, then the order of a divides $|G|$.*

The next exercise gives an important characterization of the order of an element.

EXERCISES

2.13. If $a \in G$ has finite order n, then n is the least positive integer such that $a^n = 1$. If a has infinite order, then $a^n \neq 1$ for all $n \neq 0$. (HINT: If a has finite order, define n as the least positive integer such that a^n is equal to some element on the list $1, a, a^2, \cdots, a^{n-1}$; prove that $a^n = 1$ and that $\langle a \rangle = \{1, a, a^2, \cdots, a^{n-1}\}$.)

2.14. If $|G| = 4$, then G is either cyclic or isomorphic to the group of Exercise 1.11. Conclude that a group of order 4 must be abelian.

****2.15.** Let G be a finite group with subgroups H and K. If $K \subset H \subset G$, then $[G:K] = [G:H][H:K]$.

****2.16.** If $a^n = 1$, then the order of a divides n.

2.17. If G is a group of order $2n$, then the number of elements in G of order 2 is odd. Conclude that G must contain an element of order 2.

***2.18.** If two elements a and b commute, and if $a^m = 1 = b^n$, then $(ab)^k = 1$, where $k = lcm\{m, n\}$. (The order of ab may be less than k: for example, let $b = a^{-1}$.) Conclude that if a and b have

finite order and if a and b commute, then their product ab also has finite order.

*2.19. This exercise shows that a product of two elements of finite order may have infinite order; of course this cannot happen in a finite group.

Let G be the multiplicative group of all 2×2 nonsingular matrices with rational entries. Let

$$a = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}.$$

Show that $a^4 = 1$ and $b^3 = 1$, but that ab has infinite order.

2.20. If H and K are (not necessarily distinct) subgroups of G, then a **double coset** of H, K is a subset of G of the form HtK, where $t \in G$. Prove that the family of all double cosets of H, K partitions G. (HINT: Define an equivalence relation on G by $t' \equiv t$ if $t' = htk$ for some $h \in H$ and $k \in K$.)

2.21. Suppose H, K are subgroups of a finite group G, and G is the disjoint union

$$G = \bigcup_{i=1}^{n} Ht_iK.$$

Prove that $[G:K] = \sum_{i=1}^{n} [H: H \cap t_iKt_i^{-1}]$. (Note that Lagrange's theorem is the special case of this when $K = \{1\}$.)

**2.22. Let $G = \langle a \rangle$ have order n. Prove that a^k is a generator of G if and only if $(k, n) = 1$.

Definition The **Euler φ-function** is defined as follows: If $h = 1$, then $\varphi(h) = 1$; if $h > 1$, then $\varphi(h)$ is the number of integers k such that $1 \leq k < h$ and $(k, h) = 1$.

EXERCISES

2.23. If $h = p^t$, where p is prime, then

$$\varphi(h) = p^t\left(1 - \frac{1}{p}\right).$$

2.24. Let $(r, s) = 1$. If $x \in G$ has order rs, then $x = yz$, where y has order r, z has order s, and y and z commute; the factors y and z are unique. Conclude that if $h = rs$, where $(r, s) = 1$, then

$$\varphi(rs) = \varphi(r)\varphi(s).$$

2.25. If the distinct prime divisors of h are $p_1, \cdots, p_m$, then

$$\varphi(h) = h\left(1 - \frac{1}{p_1}\right)\cdots\left(1 - \frac{1}{p_m}\right).$$

2.26. Let $U(Z_n) = \{\bar{a} \in Z_n : (a, n) = 1\}$. Prove that $U(Z_n)$ is a multiplicative group of order $\varphi(n)$.

2.27. Prove that every subgroup of a cyclic group is cyclic. (HINT: Use the division algorithm.)

***2.28.** Prove that two cyclic groups are isomorphic if and only if they have the same order.

****2.29.** Let $a \in G$ have finite order and let $f: G \to H$ be a homomorphism. Prove that the order of $f(a)$ divides the order of a.

2.30. **(Fermat)** Let p be a prime; for any integer n, show that $n^p \equiv n \pmod{p}$. (HINT: If Z_p denotes the integers modulo p, then the nonzero elements of Z_p form a multiplicative group of order $p - 1$.) Generalize to obtain **Euler's theorem**: If $(m, n) = 1$, then $n^{\varphi(m)} \equiv 1 \pmod{m}$.

2.31. **(H. B. Mann)** Let G be a finite group, and let S and T be (not necessarily distinct) nonempty subsets. Prove that either $G = ST$ or $|G| \geq |S| + |T|$.

2.32. Using the result of Mann, prove that every element in a finite field F is a sum of two squares.

NORMAL SUBGROUPS AND QUOTIENT GROUPS

There is one kind of subgroup that is especially interesting, for it is intimately related to homomorphisms.

Definition A subgroup S of G is a **normal subgroup** of G, denoted $S \lhd G$, in case $aSa^{-1} \subset S$ for every $a \in G$.

The first properties of normal subgroups are contained in the following exercises.

EXERCISES

2.33. A subgroup S of G is normal in G if and only if $aSa^{-1} = S$ for every $a \in G$.

2.34. If G is abelian, then every subgroup of G is normal. The converse is false, as the group of quaternions shows (see Exercise 5.35).

2.35. A subgroup S is normal in G if and only if every right coset of S in G is also a left coset of S in G.

2.36. Any subgroup S in G of index 2 is normal.

****2.37.** If $S \lhd G$, $a \in G$, and $s \in S$, then there exists an element $s' \in S$ with $as = s'a$ (this is a partial commutativity).

Definition Let $x \in G$. A **conjugate** of x is an element of the form axa^{-1}, where $a \in G$. One often denotes axa^{-1} by x^a.

EXERCISES

2.38. A subgroup S of G is normal if and only if, whenever $x \in S$, all conjugates of x also lie in S. Conclude that $S \lhd G$ if and only if $f(S) \subset S$ for every conjugation f (see Exercise 1.23).

2.39. Let G be the multiplicative group of all $n \times n$ nonsingular matrices over a field F. Prove that the set of all matrices of determinant 1 is a normal subgroup of G.

2.40. Any intersection of normal subgroups of G is itself a normal subgroup of G. Conclude that if X is any subset of G, there is a smallest normal subgroup of G containing X; this subgroup is called the **normal subgroup of G generated by X.**

2.41. If X is empty, the normal subgroup generated by X is $\{1\}$. If X is nonempty, the normal subgroup generated by X is the subgroup generated by all conjugates of elements in X.

****2.42.** If H and K are normal subgroups of G, then so is $H \vee K$.

The construction of a **quotient group** in the following theorem is of fundamental importance.

Theorem 2.11 *If $S \lhd G$, then the cosets of S in G form a group, denoted G/S, of order $[G:S]$.*

Proof In order to define a group, we must present a set and a binary operation. The set is the collection of all cosets of S in G; note that since $S \lhd G$, we need not bother with the adjectives "right" and "left". As multiplication, we propose the multiplication of nonempty subsets that we defined earlier. Recall that we proved then (Exercise 2.6) that this multiplication is associative. Now $(Sa)(Sb) = Sa(a^{-1}Sa)b$ (because S is normal) $= S(aa^{-1})Sab = SSab = Sab$

(because S is a subgroup). Thus, we have proved that $SaSb = Sab$, so that a product of two cosets of S is itself a coset of S. We leave to the reader the proof that the identity is S and that the inverse of Sa is Sa^{-1}, so that we have constructed a group (which is denoted G/S). Finally, the definition of index says that $|G/S| = [G:S]$. ∎

Example 1 Let m be a fixed positive integer. If k is an integer, recall that the congruence class of k modulo m is

$$\bar{k} = \{k + im: i = 0, \pm 1, \pm 2, \cdots\}.$$

The group of integers modulo m, denoted Z_m, is the set of all these congruence classes under the binary operation

$$\bar{k} + \bar{l} = \overline{k + l}.$$

If Z is the additive group of integers, then the cyclic subgroup generated by m is $Zm = \{im: i = 0, \pm 1, \pm 2, \cdots\}$. Thus $\bar{k} = k + Zm$ is a coset of Zm in Z. Now Zm is a normal subgroup of Z, since Z is abelian, so that the quotient group Z/Zm is defined. Addition in Z/Zm is given by

$$(k + Zm) + (l + Zm) = k + l + Zm.$$

Thus, Z/Zm is precisely Z_m. Because of this example, an arbitrary quotient group G/H is often called G modulo H.

Example 2 Let G be the multiplicative group of all $n \times n$ nonsingular matrices over a field F, and let H be the normal subgroup of all matrices of determinant 1. We shall show that $G/H \cong F^*$, the multiplicative group of all nonzero elements in F.

If $a \in G$, then $\det(a) \in F^*$, since a is nonsingular. Define a function $d: G/H \to F^*$ by

$$d(Ha) = \det(a).$$

This function is well defined, for if $Hb = Ha$, then $b = ha$ for some $h \in H$; therefore, $\det(h) = 1$ and

$$\det(b) = \det(ha) = \det(h)\det(a) = \det(a);$$

d is a homomorphism, for

$$d(HaHb) = d(Hab) = \det(ab) = \det(a)\det(b) = d(Ha)d(Hb);$$

d is a one-to-one, for if $d(Ha) = d(Hb)$, then $\det(a) = \det(b)$. Therefore, $\det(ab^{-1}) = 1$ and $ab^{-1} \in H$, so that $Ha = Hb$. Finally, d is onto, for if $x \in F^*$, then the diagonal matrix a that has x in the upper left corner and 1 elsewhere on the diagonal has determinant x and $d(Ha) = x$. Thus, d is an isomorphism.

THE ISOMORPHISM THEOREMS

There are three theorems that describe the relationship among quotient groups, normal subgroups, and homomorphisms. The reader should be warned that the numbering of these theorems is not canonical, so that one man's first isomorphism theorem may be another man's second. A second remark is that a testimony to the elementary character of these theorems is that analogs of them are true for every type of algebraic system, e.g., groups, rings, and vector spaces.

Theorem 2.12 (First Isomorphism Theorem) *Let $f: G \to H$ be a homomorphism with kernel K. Then K is a normal subgroup of G and $G/K \cong$ image f.*

Proof We have already seen that K is a subgroup of G. To see that K is a normal subgroup, we must show $aKa^{-1} \subset K$ for every $a \in G$. If $x = aka^{-1} \in aKa^{-1}$, then $f(x) = f(a)f(k)f(a)^{-1} = 1$ and $x \in K$, as desired.

The remainder of the proof is patterned after the example of determinants given above. Define $F: G/K \to H$ by

$$F(Ka) = f(a).$$

Now F is well defined, for if $Ka = Kb$, then $ab^{-1} \in K, f(ab^{-1}) = 1$, and so $f(a) = f(b)$. F is a homomorphism, for

$$\begin{aligned} F(KaKb) &= F(Kab) = f(ab) \\ &= f(a)f(b) = F(Ka)F(Kb). \end{aligned}$$

F is one-to-one, for if $F(Ka) = F(Kb)$, then $f(a) = f(b)$; therefore, $f(ab^{-1}) = 1$, $ab^{-1} \in K$, and so $Ka = Kb$. (That F is one-to-one is the converse of the statement that F is well defined.) Clearly, image F = image f. Therefore, F is the desired isomorphism. ∎

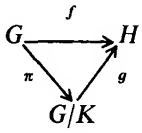
The first isomorphism theorem thus says that there is no significant difference between a quotient group and a homomorphic image.

EXERCISES

2.43. Let G be a cyclic group of order m. Use the first isomorphism theorem to prove $G \cong Z_m$. Use this result to give a new proof of Exercise 2.28.

*2.44. Let $K \lhd G$. The function $\pi: G \to G/K$ defined by $\pi(a) = Ka$ is a homomorphism onto G/K with kernel $\pi = K$ (π is called the **natural map**).

2.45. Let $f: G \to H$ be a homomorphism with kernel K. Prove that there exists a homomorphism $g: G/K \to H$ with $f = g \circ \pi$, where π is the natural map.



2.46. Let $N \lhd G$ and let $f: G \to H$ be a homomorphism whose kernel contains N. Then f induces a homomorphism $f_{\#}: G/N \to H$, namely, $f_{\#}(Na) = f(a)$.

2.47. Let $f: G \to H$ be a homomorphism; f is one-to-one if and only if kernel $f = \{1\}$.

**2.48. Let $f: G \to H$ be a homomorphism and let S be a subgroup of H. Then $f^{-1}(S)$ is a subgroup of G and contains the kernel of f.

Definition If $a, b \in G$, the **commutator** of a and b, denoted $[a, b]$, is the element $a^{-1}b^{-1}ab$. The **commutator subgroup** of G, denoted G', is the subgroup of G generated by all the commutators in G.

We shall see in Chapter 3 that a product of commutators need not be a commutator. Hence, the *set* of all commutators need not form a subgroup.

EXERCISES

2.49. Prove that G' is a normal subgroup of G. (HINT: Use Exercise 2.38.)

**2.50. If H is any normal subgroup of G, then G/H is abelian if and only if $G' \subset H$.

2.51. Let G be a finite group of odd order, and let x be the product of all the elements of G in some order. Prove that $x \in G'$.

2.52. For any group G,

$$G' = \{a_1 a_2 \cdots a_n a_1^{-1} a_2^{-1} \cdots a_n^{-1} : a_i \in G \text{ and } n \geq 2\}.$$

(HINT[1]: $(aba^{-1}b^{-1})(cdc^{-1}d^{-1})$

$$= a(ba^{-1})b^{-1}c(dc^{-1})d^{-1}a^{-1}(ab^{-1})bc^{-1}(cd^{-1})d.)$$

[1] Communicated to me by P. M. Weichsel.

With Theorem 2.12 and Exercise 2.44 we have exhibited the relationship between normal subgroups and homomorphisms: Given any homomorphism, there is a normal subgroup (its kernel); given any normal subgroup K, there is a homomorphism (the natural map) having K as its kernel. We point out that different homomorphisms can have the same kernel. For example, let $f: Z \to Z$ be the identity map and let $g: Z \to Z$ be defined by $g(n) = -n$. The kernel of each of these maps is $\{0\}$, but clearly $f \neq g$.

Theorem 2.13 *If S and T are subgroups of G and if one of them is normal, then $ST = S \vee T = TS$.*

Proof Recall that ST is just the set of products of the form st, where $s \in S$ and $t \in T$; hence, ST and TS are always subsets of $S \vee T$ that contain S and T. Thus, we need show only that ST and TS are themselves subgroups in order to show that $S \vee T$ is a subset of ST (and of TS). Let us suppose $T \lhd G$. If s_1t_1 and $s_2t_2 \in ST$, then

$$(s_1t_1)(s_2t_2)^{-1} = s_1(t_1t_2^{-1}s_2^{-1}) = s_1(s_2^{-1}t_3)$$

for some $t_3 \in T$ (this is the partial commutativity from the normality of T), and $s_1(s_2^{-1}t_3) = (s_1s_2^{-1})t_3 \in ST$. A similar proof shows that TS is a subgroup. ∎

EXERCISES

****2.53.** Let S and T be subgroups of G. Then ST is a subgroup of G if and only if $ST = TS$.

2.54. **(Modular Law)** Let A, B, and C be normal subgroups of G with $A \subset B$. If $A \cap C = B \cap C$ and $AC = BC$, then $A = B$.

***2.55.** **(Dedekind Law)** Let H, K, and L be normal subgroups of G with $H \subset L$. Then

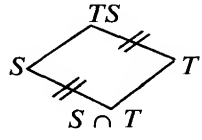
$$HK \cap L = H(K \cap L).$$

Suppose $T \subset H \subset G$ are subgroups with $T \lhd G$. Then $T \lhd H$ and the quotient group H/T is defined: It is the subgroup of G/T consisting of all cosets Th, where $h \in H$. In particular, if $T \lhd G$ and S is any subgroup of G, then $T \subset TS \subset G$, and TS/T is the subgroup of G/T consisting of all cosets Tts, where $ts \in TS$. Since $Tts = Ts$, it follows that TS/T consists precisely of all the cosets of T having a representative in S.

It follows at once from Theorem 2.13 and the product formula (Exercise 2.11) that if S and T are subgroups of G, one of which is normal, then $|TS|/|T| = |S|/|S \cap T|$. This suggests the following theorem.

Theorem 2.14 (Second Isomorphism Theorem) *Let S and T be subgroups of G with T normal. Then $S \cap T$ is normal in S, and $S/(S \cap T) \cong TS/T$.*

REMARK The accompanying diagram is a mnemonic for this theorem.



Proof Let $\pi: G \to G/T$ be the natural map and let π_0 be the restriction of π to S. Since π_0 is a homomorphism whose kernel is $S \cap T$, Theorem 2.12 tells us that $S \cap T \lhd S$ and that $S/(S \cap T) \cong$ image π_0. By our remarks above, image π_0 is just the collection of all cosets of T having representatives in S; these are precisely the cosets in TS/T. ∎

Theorem 2.15 (Third Isomorphism Theorem) *Let $K \subset H \subset G$, where both H and K are normal subgroups of G. Then H/K is a normal subgroup of G/K and*

$$(G/K)/(H/K) \cong G/H.$$

Proof Again we let the first isomorphism theorem do the dirty work. Define $f: G/K \to G/H$ by $f(Ka) = Ha$. The reader can check that f is a well-defined homomorphism whose kernel is H/K and whose image is G/H. ∎

THE CORRESPONDENCE THEOREM

The main theorem of this section could justifiably be called the fourth isomorphism theorem.

Let G and H be sets, and let $f: G \to H$ be a function. The reader is aware that f induces a "forward motion" and a "backward motion" between the subsets of G and the subsets of H. The forward motion assigns

to every subset S of G its image $f(S)$ in H; the backward motion assigns to every subset L of H its inverse image $f^{-1}(L)$ in G. Now, if f is onto, these motions define a one-to-one correspondence between all the subsets of H and some of the subsets of G. The following theorem is the group-theoretic translation of this observation.

Theorem 2.16 (Correspondence Theorem) *Let $K \lhd G$ and let $\pi: G \to G/K$ be the natural map; π defines a one-to-one correspondence between the set of those subgroups of G containing K and the set of all subgroups of G/K.*

If the subgroup of G/K corresponding to $S \subset G$ is denoted S^*, then

(i) $S^* = S/K = \pi(S)$;
(ii) *$T \subset S$ if and only if $T^* \subset S^*$, and then $[S:T] = [S^*:T^*]$;*
(iii) *$T \lhd S$ if and only if $T^* \lhd S^*$, and then $S/T \cong S^*/T^*$.*

REMARK A mnemonic diagram for this theorem is

$$\begin{array}{l} G \overset{\pi}{\to} G/K \\ S \leftrightarrow S/K \\ K \to \{1\} \end{array}$$

Proof We first show that the correspondence is one-to-one. Suppose that S and T are subgroups of G containing K and that $S/K = T/K$; we must prove that $S = T$. If $s \in S$, then $Ks = Kt$ for some $t \in T$, since $K \subset T$, and so $S \subset T$. A symmetric argument proves the reverse inclusion.

We now prove that the correspondence is onto. If A is a subgroup of G/K, we must find a subgroup S of G containing K for which $S/K = A$. By Exercise 2.48, if we define $S = \pi^{-1}(A)$, then S is a subgroup of G that contains K; furthermore, $S/K = \pi(S) = \pi(\pi^{-1}(A)) = A$, since π is onto.

It is obvious that this correspondence preserves inclusions. In order to prove $[S:T] = [S^*:T^*]$, we must show that the set of cosets Ts (where $s \in S$) is in one-to-one correspondence with the set of cosets T^*s^* (where $s^* \in S^*$). Such a correspondence is given by $Ts \leftrightarrow T^*\pi(s)$; the verification is left to the reader.[2]

If $T \lhd S$, then we conclude from the third isomorphism theorem that $T/K \lhd S/K$ and $(S/K)/(T/K) \cong S/T$, i.e., $T^* \lhd S^*$ and $S^*/T^* \cong$

[2] If all groups are finite, we may prove that $[S:T] = [S^*:T^*]$ as follows: $[S^*:T^*] = |S^*|/|T^*| = |S/K|/|T/K| = (|S|/|K|)/(|T|/|K|) = |S|/|T| = [S:T]$.

S/T. Suppose, conversely, that $T^* \lhd S^*$. Let $\eta: S^* \to S^*/T^*$ be the natural map and let $\pi_0 = \pi|S$. The reader may verify that T is the kernel of $\eta\pi_0: S \to S^*/T^*$, which implies that $T \lhd S$. ∎

EXERCISES

2.56. Let H be a subgroup of G which contains G', the commutator subgroup of G. Prove that $H \lhd G$.

****2.57.** **(Zassenhaus)** Let G be a finite group such that, for some fixed integer $n > 1$, $(xy)^n = x^n y^n$, for all $x, y \in G$. Let $G_n = \{z \in G: z^n = 1\}$ and $G^n = \{x^n: x \in G\}$. Show that both G_n and G^n are normal subgroups of G and that $|G^n| = [G:G_n]$. (NOTE: The wise reader lets homomorphisms do most of the work.)

2.58. Let G be a finite group with a normal subgroup H such that $(|H|, [G:H]) = 1$. Show that H is the unique subgroup of G having order $|H|$. (HINT: If K is another such subgroup, what happens to K in G/H?)

2.59. If $H \lhd G$, need G contain a subgroup isomorphic to G/H?

2.60. Prove that the circle group $\mathbf{T}$ is isomorphic to $\mathbf{R}/Z$, where $\mathbf{R}$ is the additive group of real numbers. (HINT: Use Exercise 1.21.)

***2.61.** Prove that H is a **maximal normal subgroup** of G (i.e., $H \lhd G$ and there is no normal subgroup strictly between G and H) if and only if G/H has no proper normal subgroups.

CHAPTER

3

Permutation Groups

PERMUTATIONS

The reader knows the formula giving the roots of a quadratic polynomial, and he is aware of similar formulas giving the roots of polynomials of degree 3 and degree 4. Many mathematicians put their faith in a faulty induction and tried to find a formula that would give the roots of an arbitrary polynomial of degree 5. In 1824, Abel proved, by studying permutations of the roots of quintic polynomials, that no such formula exists, and this result led Galois to his discovery of the intimate relationship between polynomials and certain groups of permutations of their roots. Influenced by the beauty of the work of Abel and Galois, most nineteenth-century mathematicians considered only those groups whose elements are permutations (we shall see presently that this is no restriction at all). We now proceed to develop this point of view.

Definition If X is a nonempty set, a **permutation** of X is a function $\alpha: X \to X$ that is a one-to-one correspondence.

Definition If X is a nonempty set, the **symmetric group** on X, denoted S_X, is the group whose elements are the permutations of X and whose binary operation is composition of functions.

Of particular interest is the special case when X is finite. If $X = \{1, 2, \cdots, n\}$, we write S_n instead of S_X, and we call S_n the symmetric group of degree n, or the symmetric group on n letters. Note that $|S_n| = n!$.

EXERCISES

3.1. Write a multiplication table for S_4.

****3.2.** Let X and Y be two sets and let $f: X \to Y$ be a one-to-one correspondence. Prove that $\alpha \to f \circ \alpha \circ f^{-1}$ defines an isomorphism between S_X and S_Y.

3.3. Suppose that $X \subset Y$, where X is nonempty. Prove that S_X can be imbedded in S_Y, i.e., S_X is isomorphic to a subgroup of S_Y.

Let X denote the set $\{1, 2, \cdots, n\}$. One way of denoting a permutation α of X is by displaying its values:

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha 1 & \alpha 2 & \cdots & \alpha n \end{pmatrix}.$$

Thus, $\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ and $\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ are permutations of $\{1, 2, 3\}$. Their product is $\alpha\beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ (we compute this product by first applying β and then α).[1] Note that $\beta\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, so that $\alpha\beta \neq \beta\alpha$. It follows that S_3, and hence any larger symmetric group, is not abelian.

CYCLES

In this section, we shall consider some factorizations of permutations in S_n as products of simpler permutations, called *cycles*.

Definition Let $x \in X$ and $\alpha \in S_X$; we say α **fixes** x if $\alpha(x) = x$; otherwise α **moves** x.

Definition Let $i_1, i_2, \cdots, i_r$ be distinct integers between 1 and n. If $\alpha \in S_n$ fixes the other integers and

$$\alpha(i_1) = i_2, \alpha(i_2) = i_3, \cdots, \alpha(i_{r-1}) = i_r, \alpha(i_r) = i_1,$$

then α is an **r-cycle**. We also say that α is a cycle of **length** r.

[1] Some authors compute the product $\alpha\beta$ in the reverse order.

A way to denote an r-cycle α, other than the cumbersome two-rowed notation we introduced earlier, is $\alpha = (i_1\ i_2 \cdots i_r)$. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = (2\ \ 3\ \ 4\ \ 1),$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 2 & 3 \end{pmatrix} = (1\ \ 5\ \ 3\ \ 4\ \ 2),$$

and

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} = (2\ \ 3\ \ 1).$$

All 1-cycles equal the identity permutation, which we thus denote (1).

EXERCISES

3.4. Prove that $(1\ 2\ 3 \cdots r) = (2\ 3 \cdots r1) = (3\ 4 \cdots r\ 1\ 2) = \cdots = (r\ 1\ 2 \cdots r - 1)$. Conclude that there are exactly r such notations for this cycle.

3.5. The order of an r-cycle is r.

3.6. Exhibit two 2-cycles whose product is a 3-cycle. This example shows that if α and β do not commute, nothing intelligent can be said about the order of $\alpha\beta$ in terms of the orders of the factors.

****3.7.** Let α and β be r-cycles in S_X. If there is an $x_0 \in X$ such that (i) both α and β move x_0 and (ii) $\alpha^t(x_0) = \beta^t(x_0)$ for all integers t, then $\alpha = \beta$.

Definition Two permutations α and β in S_X are **disjoint** in case every x moved by one is fixed by the other. In symbols, if $\alpha(x) \neq x$, then $\beta(x) = x$, and if $\beta(y) \neq y$, then $\alpha(y) = y$. A set of permutations is **disjoint** if each pair of them is disjoint.

EXERCISES

3.8. Let $\alpha = (i_1\ i_2 \cdots i_r)$ and $\beta = (j_1\ j_2 \cdots j_s)$. Prove that α and β are disjoint if and only if $\{i_1, i_2, \cdots, i_r\} \cap \{j_1, j_2, \cdots, j_s\} = \varnothing$.

****3.9.** If α and β are disjoint, then $\alpha\beta = \beta\alpha$.

***3.10.** A permutation $\alpha \in S_n$ is **regular** if it is (1) or if it has no fixed points and is the product of disjoint cycles of the same length. Prove that α is regular if and only if α is a power of an n-cycle.

*3.11. Let $\alpha = \beta_1\beta_2 \cdots \beta_m$, where the β_i are disjoint r_i-cycles. Prove that the order of α is $lcm\{r_1, r_2, \cdots, r_m\}$. Conclude that each r_i divides the order of α. Conclude further that if p is prime, then every power of a p-cycle is a p-cycle, or (1).

*3.12. If α is an n-cycle, then α^k is a product of (n, k) disjoint cycles, each of length $n/(n, k)$. (HINT: Use Exercises 3.10 and 3.11.)

Lemma 3.1 *If $\alpha \in S_X$, the relation on X defined by*

$$x \equiv_\alpha y \qquad \text{if } \alpha^k(x) = y \text{ for some integer } k$$

is an equivalence relation on X.

Proof Left to the reader. ∎

Definition The equivalence classes of X under the relation $x \equiv_\alpha y$ are called the **orbits** of α.

If $x_0 \in X$, the orbit of α containing x_0 is $\{y \in X: y = \alpha^k(x_0)$ for some integer $k\}$; in other words, the orbit consists of all those elements obtained from x_0 by repeatedly applying α (or α^{-1}). In order to provide examples, we now translate our earlier definitions into the language of orbits. First, "x_0 is fixed by α" means that x_0 is the sole resident of its orbit under α. Cycles provide a more interesting example. If $\alpha = (i_1\, i_2 \cdots i_r)$, then $\alpha(i_1) = i_2$, $\alpha^2(i_1) = i_3, \cdots, \alpha^{r-1}(i_1) = i_r$, so that a permutation $\neq (1)$ is a cycle if and only if it has just one orbit with more than one element, i.e., just one nontrivial orbit. A cycle is an r-cycle if its nontrivial orbit contains exactly r elements. As a final example, α is a regular permutation (Exercise 3.10) in case all its orbits have the same size.

Lemma 3.2 *If X is a finite set and $\alpha \in S_X$, then there is a partition of X into disjoint subsets $Y_1, Y_2, \cdots, Y_m$ such that, for all $i = 1, 2, \cdots, m$, $\alpha(Y_i) = Y_i$ and $\alpha|Y_i$ is a cycle.*

Proof Let $Y_1, Y_2, \cdots, Y_m$ be the orbits of α. The two properties follow immediately from the definition of $x \equiv_\alpha y$. ∎

Theorem 3.3 *Every permutation $\alpha \neq (1)$ in S_n is the product of disjoint cycles of length ≥ 2; this factorization is unique except for the order in which the cycles are written.*

Proof Let $Y_1, Y_2, \cdots, Y_m$ be the orbits of α. Lemma 3.2 tells us that each $\alpha|Y_i$ is a cycle. If we define $\beta_i \in S_n$ $(i = 1, 2, \cdots, m)$ as that permutation which is α on Y_i and which fixes everything else, then

β_i is a cycle in S_n. Further, the β_i are pairwise disjoint and $\alpha = \beta_1\beta_2 \cdots \beta_m$ (just evaluate both sides). We obtain the desired factorization if we delete all factors β_i equal to the identity.

We now prove the uniqueness assertion. Suppose

$$\alpha = \beta_1\beta_2 \cdots \beta_m = \gamma_1\gamma_2 \cdots \gamma_k,$$

where each of these is a factorization of α into pairwise disjoint cycles of length ≥ 2. If α moves x, then one of the β and one of the γ must also move x. Since disjoint cycles commute, we may assume for notational convenience that β_1 and γ_1 move x. For any integer t,

$$\beta_1^t(x) = \alpha^t(x) = \gamma_1^t(x).$$

Since, by Exercise 3.7, a cycle is completely determined by the behavior of its powers on any element it moves, $\beta_1 = \gamma_1$. We now cancel these factors, and the proof is completed by an induction on $\max\{m, k\}$. ∎

EXERCISES

3.13. Write the following permutation as a product of disjoint cycles and compute its order:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 4 & 7 & 2 & 5 & 1 & 8 & 9 & 3 \end{pmatrix}.$$

****3.14.** If p is a prime, prove that the only elements of order p in S_n are products of disjoint p-cycles.

Definition A 2-cycle is also called **transposition**.

Of all the permutations, surely the transposition, which merely interchanges two points, is the simplest.

Theorem 3.4 *Every $\alpha \in S_n$ is a product of transpositions.*

Proof By Theorem 3.3 it suffices to factor any r-cycle into a product of transpositions. This is done in the following way:

$$(1\ 2\ \cdots\ r) = (1\ r)(1\ \ r-1)\cdots(1\ 2). \quad ∎$$

Any permutation can thus be realized as a sequence of interchanges. This factorization, however, is not as nice as the factorization into disjoint cycles. First of all, the transpositions occurring need not commute; e.g., $(123) = (13)(12) \neq (12)(13)$. Second, the factors are not uniquely

determined; e.g., (123) = (13)(12) = (23)(12) = (13)(42)(12)(14) = (13)(42)(12)(14)(23)(23). Is there any uniqueness in such a factorization? We shall prove that the number of factors is always even or always odd.

Consider the polynomial in n variables:

$$g(x_1, x_2, \cdots, x_n) = \prod_{j>i} (x_j - x_i).$$

Notice that the condition $j > i$ means that every pair of distinct integers between 1 and n occurs as subscripts in exactly one factor of g. If $\alpha \in S_n$, we form a new polynomial in the n variables:

$$\alpha g = \prod_{j>i} (x_{\alpha(j)} - x_{\alpha(i)}).$$

Since α is a permutation, every pair of distinct integers between 1 and n occurs in exactly one factor of αg. Therefore, except for signs, the factors of g and αg are identical. Hence, $\alpha g = \pm g$.

Definition A permutation $\alpha \in S_n$ is **even** if $\alpha g = g$; α is **odd** if $\alpha g = -g$.

If $\alpha, \beta \in S_n$, then $(\alpha\beta)g = \alpha(\beta g)$, for both are equal to

$$\prod_{j>i} (x_{\alpha\beta(j)} - x_{\alpha\beta(i)}).$$

It follows that if both α and β are even or if both α and β are odd, then $\alpha\beta$ is even, while if one factor is even and the other odd, then $\alpha\beta$ is odd.

Definition The **alternating group** of degree n, denoted A_n, is the set of all even permutations in S_n.

Define a function $\varphi: S_n \to G$, where G is the multiplicative group $\{1, -1\}$, by $\varphi(\alpha) = 1$ if α is even and $\varphi(\alpha) = -1$ if α is odd. Our preceding remarks prove that φ is a homomorphism. Given a homomorphism, we ask the canonical questions: What is kernel φ? What is image φ?

Theorem 3.5 *A_n is a normal subgroup of S_n of order $\frac{1}{2}n!$.*

Proof It follows from the definitions of A_n and the homomorphism φ that A_n = kernel φ. We claim that image $\varphi = \{1, -1\}$, i.e., that φ is onto. In order to prove this, it suffices to exhibit an odd permutation. The transposition $\tau = (1\ 2)$ is odd, for the only factor in $g(x_1, x_2, \cdots, x_n)$ that changes sign is $x_2 - x_1$. By the first isomor-

phism theorem, $A_n \lhd S_n$ and $S_n/A_n \cong \{1, -1\}$. Therefore, $|A_n| = \frac{1}{2}n!$. ∎

Theorem 3.6 *$\alpha \in S_n$ is even if and only if α is a product of an even number of transpositions; $\alpha \in S_n$ is odd if and only if α is a product of an odd number of transpositions.*

Proof We first prove that every transposition $\tau = (i\,j)$ is odd. This could be proved by actually considering the action of τ on $g(x_1, x_2, \cdots, x_n)$, but that would be tedious. Suppose, instead, that $(i\,j)$ is even, i.e., $(i\,j) \in A_n$. Since $A_n \lhd S_n$, every conjugate of $(i\,j)$ is also in A_n. But (1 2) is a conjugate of $(i\,j)$: if $\gamma = (i\ 1)(j\ 2)$, then $\gamma(i\,j)\gamma^{-1} = (i\ 1)(j\ 2)(i\,j)(j\ 2)(i\ 1) = (1\ 2)$, contradicting the fact that (1 2) is odd.[2]

Suppose $\alpha \in S_n$ and $\alpha = \tau_1\tau_2 \cdots \tau_m$, where each τ_i is a transposition. Then $\varphi(\alpha) = \varphi(\tau_1\tau_2 \cdots \tau_m) = \varphi(\tau_1)\varphi(\tau_2) \cdots \varphi(\tau_m) = (-1)^m$. Now α is even if and only if $\varphi(\alpha) = 1$, and α is odd if and only if $\varphi(\alpha) = -1$. Therefore, α is even if and only if m is even; and α is odd if and only if m is odd. ∎

Corollary 3.7 *If $\alpha \in S_n$, the number of factors occurring in any factorization of α into transpositions is always even or always odd.*

Proof A permutation cannot be both even and odd. ∎

EXERCISES

****3.15.** An $n \times n$ **permutation matrix** is a matrix obtained from the $n \times n$ identity matrix E by permuting its columns. If P_n is the set of all $n \times n$ permutation matrices, prove that P_n is a multiplicative group and that $\theta: S_n \to P_n$ is an isomorphism, where $\theta(\alpha)$ is the matrix obtained from E by permuting its columns according to α. Prove that α is even (or odd) if and only if $\det\theta(\alpha)$ is 1 (or -1).

3.16. An r-cycle is even if and only if r is odd.

****3.17.** If $n > 2$, then A_n is generated by the 3-cycles.

3.18. If a subgroup G of S_n contains an odd permutation, then $|G|$ is even and exactly half the elements of G are odd permutations.

[2] This computation is a special case of a theorem we shall soon prove which tells precisely when two permutations are conjugate in S_n. We must also be a bit more careful: If $i = 1$, define $\gamma = (j\ 2)$; if $j = 2$, define $\gamma = (i\ 1)$.

3.19. Imbed S_n as a subgroup in A_{n+2}. Conclude that **parity** (being even or odd) is not preserved by homomorphisms.

****3.20.** Prove that S_n can be generated by (12), (13), $\cdots$, (1n); prove that S_n can also be generated by (1 2) and (12$\cdots n$).

CONJUGATES

In this section we study conjugates and conjugacy classes, first for arbitrary groups and then for the special case of symmetric groups.

We begin by defining another equivalence relation.

Definition If G is a group, the relation "x is a conjugate of y in G" is an equivalence relation on G; the equivalence classes are called **conjugacy classes.**

As an example, if G is the multiplicative group of all $n \times n$ nonsingular matrices over a field F, then two matrices lie in the same conjugacy class if and only if they are similar.

Now x and y lie in the same conjugacy class if there is an element $a \in G$ with $y = axa^{-1} = x^a$. There is thus an isomorphism $f: G \to G$ (namely, conjugation by a) with $y = f(x)$. It follows that all elements in the same conjugacy class have the same order. An amusing consequence is that, for any two elements a and b in G, ab and ba have the same order.

An element $x \in G$ is the sole resident of its conjugacy class if $x = axa^{-1}$ for all $a \in G$, i.e., x commutes with every element in G. In an abelian group, therefore, conjugacy classes are not of much interest.

Definition The **center** of G, denoted $Z(G)$, is the set of all $x \in G$ that commute with every element in G.

EXERCISES

3.21. $Z(G)$ is a normal, abelian subgroup of G.

****3.22.** If G is a nonabelian group, then $G/Z(G)$ is not cyclic.

****3.23.** A subgroup S of G is a normal subgroup of G if and only if S is a (set theoretical) union of conjugacy classes of G.

***3.24.** Let G be a group containing an element of finite order $n > 1$ and exactly two conjugacy classes. Prove that G is a cyclic group of order 2.

It is very useful to count the number of elements in a conjugacy class. To this end, we introduce the following subgroup.

Definition If $x \in G$, the **centralizer** of x in G, denoted $C_G(x)$, is the set of all $a \in G$ that commute with x.

It is immediate that $C_G(x)$ is a subgroup of G. When the meaning is clear from the context, we shall abbreviate $C_G(x)$ by $C(x)$.

Theorem 3.8 *The number of conjugates of x in G is $[G: C_G(x)]$, and hence this number is a divisor of $|G|$ when G is finite.*

Proof If a and b are elements of G, then the following statements are equivalent:

$axa^{-1} = bxb^{-1}$;
$a^{-1}b$ commutes with x;
$a^{-1}b \in C(x)$;
a and b lie in the same left coset of $C(x)$, i.e., $aC(x) = bC(x)$.

The function ψ, defined by $\psi(axa^{-1}) = aC(x)$, is thus a one-to-one correspondence between the set of distinct conjugates of x and the left cosets of $C(x)$. ∎

We return to symmetric groups, and we ask when two permutations are conjugate in S_n.

Lemma 3.9 *If α and β are in S_n, then $\alpha\beta\alpha^{-1}$ is the permutation that has the same cycle structure as β and that is obtained by applying α to the symbols in β.*

Example If $\beta = (13)(247)$ and $\alpha = (256)(143)$, then $\alpha\beta\alpha^{-1} = (\alpha 1\ \alpha 3)(\alpha 2\ \alpha 4\ \alpha 7) = (41)(537)$.

Proof Let $\beta = \gamma_1\gamma_2 \cdots (\cdots ij \cdots) \cdots \gamma_t$ be a factorization of β into disjoint cycles. In order to prove that two permutations are equal, we must show that they have the same effect on each symbol. We have pictured β so that $\beta(i) = j$. Suppose $\alpha(i) = k$ and $\alpha(j) = m$. The instructions in the lemma say that we should send k into m. On the other hand, $\alpha\beta\alpha^{-1}$ sends $k \to i \to j \to m$, so that both permutations have the same effect on k. Since $k = \alpha(i)$ is a typical integer between 1 and n (α is onto), the two permutations are equal. ∎

Theorem 3.10 *α and β in S_n are conjugate in S_n if and only if they have the same cycle structure.*

Proof Our lemma shows that conjugate permutations have the same cycle structure.

Conversely, let α and β have the same cycle structure. Define $\gamma \in S_n$ as follows: Place α over β so that the cycles of the same length correspond; let γ: top $\rightarrow$ bottom. By the lemma, $\gamma\alpha\gamma^{-1} = \beta$, so that α and β are conjugate. ∎

For example, if

$$\alpha = (231)(45)(6)$$

and

$$\beta = (462)(31)(5),$$

then $\gamma\alpha\gamma^{-1} = \beta$, where

$$\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 3 & 1 & 5 \end{pmatrix} = (1 \quad 2 \quad 4 \quad 3 \quad 6 \quad 5).$$

If $1 < k \leq n$, then there are $(1/k)[n(n-1)\cdots(n-k+1)]$ distinct k-cycles in S_n. This formula may be used to compute the number of permutations α in S_n of a given cycle structure if one is careful about the case when several factors of α have the same length. For example, the number of permutations in S_4 of the form $(ab)(cd)$ is

$$\frac{1}{2}\left(\frac{4 \times 3}{2} \times \frac{2 \times 1}{2}\right),$$

the factor $1/2$ occurring in order that we not count $(ab)(cd) = (cd)(ab)$ twice.

Let us now examine S_4 in some detail, using the accompanying table.

Cycle Structure	Number	Order	Parity
(1)	1	1	Even
(12)	$6 = (4 \times 3)/2$	2	Odd
(123)	$8 = (4 \times 3 \times 2)/3$	3	Even
(1234)	$6 = 4!/4$	4	Odd
(12)(34)	$3 = \frac{1}{2}\left(\frac{4 \times 3}{2} \times \frac{2 \times 1}{2}\right)$	2	Even
	$24 = 4!$		

Thus, the 12 elements of A_4 are eight 3-cycles, three products of disjoint transpositions, and the identity. These elements are: (123), (132), (234), (243), (341), (314), (412), (421), (14)(23), (12)(34), (13)(24), and the identity (1).

We now show that the converse of Lagrange's theorem is false.

Theorem 3.11 *A_4 is a group of order 12 having no subgroup of order 6.*

Proof If H is such a subgroup, then $[A_4:H] = 2$, so that H is a normal subgroup. Thus, if $\alpha \in H$, H contains all conjugates of α by *even* permutations.

Since 8 of the 12 elements of A_4 are 3-cycles, H must contain a 3-cycle α. How many conjugates does α have in H? By Theorem 3.10, α has 8 conjugates in S_4, so that the centralizer of α in S_4 has index 8 and order $24/8 = 3$. There are thus only 3 permutations in S_4 (a fortiori, in the smaller A_4) that commute with α; these must be α, α^{-1}, and (1). Since these all lie in A_4, the centralizer of α in A_4 has order 3 and hence has index $12/3 = 4$. Therefore, α has 4 conjugates in A_4, and all lie in H, since H is normal.

Now H must contain an element β of order 2, for if it contained a fifth 3-cycle, it would also contain its 4 conjugates, and there would be more than 6 elements in H. Since β is even, $\beta = (ab)(cd)$. We have discovered 6 elements in H: 4 3-cycles, β, and (1), and this is our quota. But we can exhibit another element in H! Since $H \lhd A_4$, $\gamma\beta\gamma^{-1} \in H$, where $\gamma = (acb)$; but $\gamma\beta\gamma^{-1} = (ca)(bd) \neq \beta$. This contradiction completes the proof. ∎

It is now easy to give an example showing that strict inequality may hold in Exercise 2.11 when neither subgroup is normal. Let S and T be cyclic subgroups of A_4 of orders 2 and 3, respectively. Then $|S\|T| = 6$, but $|S \cap T\|S \vee T| = |S \vee T| = 12$, since A_4 has no subgroup of order 6.

EXERCISES

****3.25.** Prove that A_4 is the only subgroup of S_4 having order 12.

****3.26.** Let G be a finite group with subgroup H of index 2. If $x \in H$ has m conjugates in G, then x has either m or $m/2$ conjugates in H.

3.27. (Carmichael) Let G be the subgroup of S_{16} generated by the following eight elements:

$$(ac)(bd);\ (eg)(fh);$$
$$(ik)(jl);\ (mo)(np);$$
$$(ac)(eg)(ik);\ (ab)(cd)(mo);$$
$$(ef)(gh)(mn)(op);\ (ij)(kl).$$

Prove that G' is generated by the first four elements listed above, and that $|G'| = 16$. Moreover,

$$\alpha = (ik)(jl)(mo)(np) \in G',$$

but α is not a commutator.

THE SIMPLICITY OF A_n

Definition A group G is **simple** if $G \neq \{1\}$ and it contains no proper normal subgroups.

EXERCISES

3.28. If $H \lhd G$, prove that G/H is simple if and only if H is a maximal normal subgroup of G (see Exercise 2.61).

3.29. Prove that an abelian group is simple if and only if it is finite and of prime order.

3.30. (Schur) Let $f\colon G \to H$ be a nontrivial homomorphism, i.e., f does not send every element into 1. If G is simple, then f is one-to-one.

3.31. Prove that no group of order 12 is simple. Conclude that A_4 is not simple.

Our immediate goal is to provide the reader with some examples of simple groups. The classification of all finite simple groups is still an unsolved problem, but there are many simple groups known other than the ones we shall exhibit.

We shall prove that A_n is simple when $n > 4$. Now the subset of A_4

$$V = \{(1),\ (12)(34),\ (13)(24),\ (14)(23)\}$$

is easily seen to be a subgroup of A_4 (which is called the **Klein 4-group**). Since V contains all the permutations in S_4 of a given cycle structure, V is a normal subgroup of S_4, a fortiori, $V \lhd A_4$. Therefore, A_4 is not simple. Let us now examine S_5 and A_5.

Cycle Structure	S_5 Number	Order	Parity
(1)	1	1	Even
(12)	$10 = (5 \times 4)/2$	2	Odd
(123)	$20 = (5 \times 4 \times 3)/3$	3	Even
(1234)	$30 = (5 \times 4 \times 3 \times 2)/4$	4	Odd
(12345)	$24 = 5!/5$	5	Even
(12)(34)	$15 = \frac{1}{2}\left(\frac{5 \times 4}{2} \times \frac{3 \times 2}{2}\right)$	2	Even
(123)(45)	$20 = \frac{5 \times 4 \times 3}{3} \times \frac{2 \times 1}{2}$	6	Odd
	$120 = 5!$		

Cycle Structure	A_5 Number	Order	Parity
(1)	1	1	Even
(123)	20	3	Even
(12345)	24	5	Even
(12)(34)	15	2	Even
	60		

Lemma 3.12 *A_5 is simple.*

Proof

(i) *All 3-cycles are conjugate in A_5.* (We know this is true in S_5, but now we are only allowed to conjugate by even permutations.)

In S_5 a 3-cycle α has 20 conjugates (for our table shows that there are 20 3-cycles). Hence, $C_S(\alpha)$ has index 20 and order 6. We can exhibit those 6 elements that commute with α: if, for example, $\alpha = (1\ 2\ 3)$, then the elements of $C_S(\alpha)$ are (1), (123), (132), (45), (123)(45), (132)(45). Now only the first three of these are even, so that $C_A(\alpha)$ has order 3 and hence index 20 in A_5. Therefore, α has 20 conjugates in A_5, so that all 3-cycles are conjugate in A_5.

(ii) *All products of disjoint transpositions are conjugate in A_5.*

If, for example, $\alpha = (12)(34)$, then our table shows that α has 15 conjugates in S_5. By Exercise 3.26, α has either 15 or 15/2 conjugates in A_5, and the latter is clearly impossible.

(iii) *There are two conjugacy classes of 5-cycles in A_5, each of which has 12 elements.*

In S_5, $\alpha = (12345)$ has 24 conjugates, so that $C_S(\alpha)$ has 5 elements, and these must be the powers of α. Therefore, $|C_A(\alpha)| = 5$, and so the number of conjugates in A_5 is $60/5 = 12$.

We have now surveyed all the conjugacy classes occurring in A_5. If H is a normal subgroup of $A_5 \neq \{1\}$, then H is a union of conjugacy classes of A_5. The order of H is thus a sum of certain of the numbers 1, 20, 15, 12, 12. Since H contains (1), it is easily checked that one never gets a proper divisor of 60. Therefore, $|H| = 60$ and $H = A_5$; it follows that A_5 is simple. ∎

Lemma 3.13 *Let $H \lhd A_n$, where $n \geq 5$. If H contains a 3-cycle, then $H = A_n$.*

Proof We shall first show that (123) and $(i\,j\,k)$ are conjugate in A_n (and thus that all 3-cycles are conjugate in A_n). If these cycles are not disjoint, at most 5 symbols are moved by them. Let $A^* \subset A_n$ be the alternating group on these symbols (if the cycles move less than 5 symbols, just supplement them to get 5 symbols). As in part (i) of the proof of Lemma 3.12, (123) and $(i\,j\,k)$ are conjugate in A^*; a fortiori, they are conjugate in A_n. If the cycles are disjoint, then we have just seen that (123) is conjugate to $(i\,2\,3)$ and $(i\,2\,3)$ is conjugate to $(i\,j\,k)$; therefore, (123) is conjugate to $(i\,j\,k)$.

Since H is normal in A_n, it must contain all 3-cycles once it contains one of them. But A_n is generated by the 3-cycles; therefore, $H = A_n$. ∎

Theorem 3.14 *A_n is simple for $n \geq 5$.*

Proof Let $H \neq \{1\}$ be a normal subgroup of A_n, where $n \geq 5$. If $\beta \in H$ and $\beta \neq (1)$, there is a symbol i moved by β: $\beta(i) = j \neq i$. If α is any 3-cycle that fixes i and moves j, then α and β do not commute: $\beta\alpha(i) = \beta(i) = j$ and $\alpha\beta(i) = \alpha(j) \neq j$. Therefore the commutator $\alpha\beta\alpha^{-1}\beta^{-1} \neq (1)$; moreover, $(\alpha\beta\alpha^{-1})\beta^{-1}$ lies in the normal subgroup H and is a product of two 3-cycles $\alpha(\beta\alpha^{-1}\beta^{-1})$. Thus, $\alpha(\beta\alpha^{-1}\beta^{-1})$ moves at most 6 symbols.

If $\alpha(\beta\alpha^{-1}\beta^{-1})$ moves at most 5 symbols, $i_1, \cdots, i_5$, then $\alpha(\beta\alpha^{-1}\beta^{-1}) \in H \cap A_5^*$, where A_5^* is the subgroup of A_n fixing all symbols $j \neq i_1, \cdots, i_5$. Since $A_5^* \cong A_5$ is a simple group and $\{1\} \neq H \cap A_5^* \lhd A_5^*$, it follows that $H \cap A_5^* = A_5^*$, i.e., $A_5^* \subset H$.

Therefore H contains a 3-cycle, so that $H = A_n$, by Lemma 3.13, and A_n is simple.

To force $\alpha(\beta\alpha^{-1}\beta^{-1})$ to move at most 5 symbols, let us choose α more carefully. So far, α is any 3-cycle that fixes i and moves j, where $\beta(i) = j \neq i$. Assume $\beta(j) = k \neq i$. Define $\alpha = (j\ k\ l)$, where $l \neq i$. Then $\alpha^{-1} = (j\ l\ k)$ and

$$\begin{aligned}\alpha(\beta\alpha^{-1}\beta^{-1}) &= (j\ k\ l)(\beta j\ \beta l\ \beta k)\\ &= (j\ k\ l)(k * *),\end{aligned}$$

a permutation moving at most 5 symbols. Finally, assume $\beta(j) = i$. By Theorem 3.3, $\beta = (i\,j)\gamma$, where $(i\,j)$ and γ are disjoint permutations. Moreover, $\gamma \neq (1)$, since β is an even permutation, so there is a symbol $k \neq i, j$ with $\beta(k) = \gamma(k) = l \neq k$. Note that $l \neq i$ since $(i\,j)$ and γ are disjoint. If we define $\alpha = (j\ l\ k)$, then α fixes i and

$$\begin{aligned}\alpha(\beta\alpha^{-1}\beta^{-1}) &= (j\ k\ l)(\beta j\ \beta l\ \beta k)\\ &= (j\ k\ l)(i * l),\end{aligned}$$

a permutation moving at most 5 symbols. ∎

EXERCISES

***3.32.** If $n \neq 4$, A_n is the only proper normal subgroup of S_n.

***3.33.** Consider the chain of subgroups:

$$S_4 \supset A_4 \supset V \supset W \supset \{(1)\},$$

where $W = \langle(12)(34)\rangle$. We have already seen that V (the 4-group) is normal in A_4. Show that W is normal in V, but that W is not normal in A_4. Conclude that normality need not be transitive.

****3.34.** A group G is called **centerless** if $Z(G) = \{1\}$. Prove that S_n is centerless for $n \geq 3$ and that A_n is centerless for $n \geq 4$.

3.35. Give an example of a group G with more than one element such that $G = G'$ (G' is the commutator subgroup of G).

3.36. The following is an example of an infinite simple group. Let X be the set of positive integers and let S be the group of all permutations of X; let F be the subgroup of S consisting of all α that move only finitely many elements of X.

The subgroup of F generated by all 3-cycles, denoted A_∞, is the **infinite alternating group**. Prove that A_∞ is an infinite simple group. (HINT: Use the proof of Theorem 3.14.)

SOME REPRESENTATION THEOREMS

A very useful technique in studying a group is to represent it in terms of something familiar and concrete. If, for example, the elements of a group happen to be permutations or matrices, we may be able to obtain results by using this extra information. In this section, we give some elementary theorems on representations, i.e., on homomorphisms into familiar groups.

The following theorem shows that the study of groups of permutations is no less general than the study of arbitrary groups.

Theorem 3.15 (Cayley) *Every group G is isomorphic to a subgroup of S_G. In particular, every finite group of order n is isomorphic to a subgroup of S_n.*

Proof For each $a \in G$, define $T_a: G \to G$ by $T_a(x) = ax$ (T_a is left translation by a). In Exercise 1.5, we saw that T_a is a one-to-one correspondence, so that $T_a \in S_G$. The function τ sending a into T_a is thus a function from G to S_G.

We claim that τ is one-to-one and a homomorphism. If $a \neq b$ are elements of G, then $T_a(1) = a \neq b = T_b(1)$, so that $T_a \neq T_b$ and $\tau(a) \neq \tau(b)$; therefore, τ is one-to-one. Finally, consider $\tau(ab) = T_{ab}$ and $\tau(a)\tau(b) = T_aT_b$. In order to show that these permutations are the same, we must show that they assign the same value to each $x \in G$. But $T_{ab}(x) = (ab)x$, and $T_aT_b(x) = T_a(bx) = a(bx)$, and these are the same by associativity. ∎

EXERCISES

3.37. Prove that A_∞ contains an isomorphic copy of every finite group.

3.38. If G is a finite group, each T_a in the above proof is a regular permutation. (HINT: The orbits of T_a are the right cosets of $\langle a \rangle$ in G.) For this reason, the homomorphism τ is called the (left) **regular representation** of G.

3.39. Why doesn't the proof of Cayley's theorem show that every cancellation semigroup can be imbedded in a group?

****3.40.** Prove that every finite group can be imbedded in a group that can be generated by two elements.

3.41. Let G be a group of order $2^m k$, where k is odd. If G contains an element of order 2^m, then the set of all elements in G having odd order is a (normal) subgroup of G. (HINT: Consider G as permutations via Cayley's theorem and show that it contains an odd

permutation.) Conclude that a finite nonabelian simple group of even order must have order divisible by 4.

Corollary 3.16 *Let G be a finite group of order n and let F be a field. Then G can be imbedded in $GL(n, F)$, the multiplicative group of all nonsingular $n \times n$ matrices with entries in F.*[3]

Proof By Cayley's theorem, there is an imbedding of G into S_n. By Exercise 3.15, there is an isomorphism $S_n \cong P_n$, the multiplicative group of all $n \times n$ permutation matrices. Since the entries of a permutation matrix are only 0 and 1, and since it is only the multiplicative properties of 0 and 1 that are needed in establishing the isomorphism of Exercise 3.15, we may regard 0 and 1 as lying in the field F. ∎

Theorem 3.17 *Let B be a subgroup of index n in a group G. There is a homomorphism $\rho: G \to S_n$ whose kernel is contained in B. Indeed, kernel $\rho = \bigcap_{x \in G} xBx^{-1}$.*

Proof Let $X = \{g_1B, \cdots, g_nB\}$ be the family of all left cosets of B in G. If $a \in G$, define a function $\rho_a: X \to X$ by $g_iB \to ag_iB$, for all i. It is easy to check that each ρ_a is a permutation and that $\rho: G \to S_X \cong S_n$ defined by $a \to \rho_a$ is a homomorphism.

If $a \in \ker \rho$, then ρ_a is the identity permutation, i.e., $ag_iB = g_iB$ for all i. Thus there are elements b_i, b_i' in B with $ag_ib_i = g_ib_i'$ for all i. Therefore,

$$a = g_ib_i'b_i^{-1}g_i^{-1} \in g_iBg_i^{-1} \qquad \text{for all } i,$$

and so

$$a \in \bigcap_i g_iBg_i^{-1}.$$

Since every subgroup $xBx^{-1} = g_iBg_i^{-1}$ for some i, it follows that kernel $\rho \subset \bigcap_{x \in G} xBx^{-1}$. The reverse inclusion is easy to prove and is left to the reader. ∎

It follows immediately that $\bigcap_{x \in G} xBx^{-1}$ is a normal subgroup of G, but this also follows easily from the definitions.

Corollary 3.18 *If an infinite group G contains a proper subgroup B of finite index, then G contains a proper normal subgroup of finite index.*

[3] *GL* abbreviates "general linear". There is also a "special linear" group that we shall discuss in Chapter 8.

Proof Let B have index n. By Theorem 3.17, there is a homomorphism $\rho: G \to S_n$ whose kernel K is contained in B. Therefore, $K \lhd G$, and since G/K is isomorphic to a subgroup of the finite group S_n, K has finite index in G. ∎

Corollary 3.19 *A simple group G containing a proper subgroup B of index n can be imbedded in S_n.*

Proof By Theorem 3.17, there is a homomorphism $\rho: G \to S_n$ whose kernel K is contained in B. Since G is simple, $K = \{1\}$ or $K = G$; since B is proper, $K \subset B \neq G$. Therefore, $K = \{1\}$ and ρ is one-to-one. ∎

Corollary 3.19 provides a substantial improvement over Cayley's theorem, at least for simple groups. For example, if $G \cong A_5$, then Cayley's theorem asserts that G can be imbedded in S_{60}. On the other hand, G contains a subgroup B of order 12 (isomorphic to A_4) and hence of index 5; therefore, Corollary 3.19 asserts that G can be imbedded in S_5.

EXERCISES

3.42. **(Poincaré)** If H and K are subgroups of G having finite index, then $H \cap K$ has finite index in G.

3.43. Let G be a finite group containing a subgroup H with $[G:H] = p$, where p is the smallest prime divisor of $|G|$. Prove that H is normal in G.

3.44. Let H be a proper subgroup of G. Prove there is a group K and distinct homomorphisms $f, g: G \to K$ with $f|H = g|H$. (HINT: Let X be the family of all left cosets of H in G together with an additional element denoted ∞, and let $K = S_X$. Define $f: G \to S_X$ as follows: If $t \in G$, then $f_t(\infty) = \infty$ and $f_t(aH) = taH$. Define $g: G \to S_X$ by $g = \beta \circ f$, where $\beta: S_X \to S_X$ is conjugation by the transposition that interchanges H and ∞.)

3.45. A homomorphism $h: A \to G$ is onto if and only if for every group K and every pair of homomorphisms $f, g: G \to K$, $f \circ h = g \circ h$ implies $f = g$.

COUNTING ORBITS

Let G be a group of permutations acting on a set X, i.e., G is a subgroup of S_X. Two elements x and y in X are **G-equivalent** in case there is some $t \in G$ with $t(x) = y$.

Definition The G-equivalence classes of X are called the **orbits** of G.

If $G = \langle \alpha \rangle$, where α is a permutation, then this notion of equivalence and orbit is precisely our earlier one. If G is not cyclic, we have a generalization of our previous definition, which allows us to translate several earlier definitions into the language of orbits. By the Cayley theorem, any subgroup S of G is a group of permutations on G: If $s \in S$, s acts on G by left translation. The orbits of S are just the right cosets of S in G, as the reader should verify.

A second example is provided by a group G acting as a group of permutations on itself via conjugations. In this case, the orbits are just the conjugacy classes of G.

Definition Let G be a group of permutations on a set X and let $x \in X$. The **stabilizer** of x, H_x, is the subgroup of G

$$H_x = \{t \in G: t \text{ fixes } x\}.$$

Theorem 3.20 *Let G be a group of permutations on a finite set X and let $x \in X$. Then the size of the orbit of x is $[G:H_x]$.*

Proof Let $t_1H_x, \cdots, t_rH_x$ be the distinct left cosets of H_x in G. If $t \in t_iH_x$, then $t = t_ih$ for some $h \in H_x$ and $t(x) = t_i(h(x)) = t_i(x)$; if $i \neq j$, then $t_i(x) \neq t_j(x)$ lest $t_i^{-1}t_j$ fix x and $t_i^{-1}t_j \in H_x$. The size of the orbit of x is thus $r = [G:H_x]$. ∎

Corollary 3.21 *If G is a group of permutations on a finite set X, then the size of every orbit of G divides $|G|$.*

Corollary 3.22 *Assume G is finite and $\alpha: G \to S_n$ is a homomorphism. Then the size of every orbit of $\alpha(G)$ (considered as permutations on n letters) divides $|G|$.*

Theorem 3.23 *Let X be a finite set and let G be a group of permutations on X. If N is the number of orbits of G, then*

$$N = \frac{1}{|G|} \sum_{t \in G} F(t),$$

where $F(t)$ is the number of x in X that are fixed by t.

Proof In the sum $\sum_{t \in G} F(t)$, each $x \in X$ is counted $|H_x|$ times. If x and y lie in the same orbit, then Theorem 3.20 implies that $|H_x| = |H_y|$, for $[G:H_x] = [G:H_y]$. Therefore, the $[G:H_x]$ elements constituting the orbit of x are, in the above sum, collectively counted

$[G:H_x]\,|H_x| = |G|$ times. Each orbit thus contributes $|G|$ to the sum, so that $\sum_{t\in G} F(t) = N|G|$. ∎

EXERCISES

3.46. Let G be a group of permutations on a set X and let $x, y \in X$. If $t(x) = y$ for some $t \in G$, then H_x and H_y are isomorphic.

3.47. Use Theorem 3.23 to prove Lagrange's theorem.

3.48. In the case of G acting on itself by conjugation, prove that the stabilizer of $x \in G$ is $C_G(x)$, the centralizer of x.

3.49. If G is a finite group and c is the number of conjugacy classes in G, then

$$c = \frac{1}{|G|} \sum_{x\in G} |C_G(x)|.$$

It is clear that we get more information about a permutation group G if we know it lies in a small S_n rather than in a larger one. One indication that the imbedding of G in S_n is best (i.e., that n is as small as possible) is that G is "dense" in S_n in the sense of the following definition.

Definition A permutation group G on X is **transitive** in case, for each x and $y \in X$, there is a $t \in G$ with $t(x) = y$.

(One also speaks of degrees of transitivity, which indicate that a group is even more densely imbedded in a symmetric group.)

EXERCISES

3.50. Let G be finite and let $n < |G|$. G is isomorphic to a transitive subgroup of S_n if and only if G contains a subgroup H of index n such that neither H nor any proper subgroup of H is normal in G. (HINT: For sufficiency, use Theorem 3.17; for necessity, take for H the stabilizer of any symbol.)

***3.51.** If G is a transitive subgroup of S_n, then n divides $|G|$.

We may now give a swift proof of Exercise 3.24, at least when G is finite. If G has only 2 conjugacy classes, then every two nonidentity elements of G are conjugate. Thus, G acts transitively (via conjugation) on $G - \{1\}$. Hence, Exercise 3.51 says $|G| - 1$ divides $|G|$, so that $|G| = 2$.

Definition Let G be a group of permutations on a set X. G is **doubly**

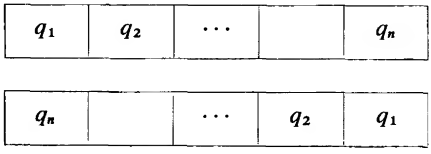
transitive if, for every two ordered pairs x_1, x_2 and y_1, y_2 of distinct elements in X, there is an element $t \in G$ with $t(x_1) = y_1$ and $t(x_2) = y_2$.

EXERCISES

3.52. If G is a doubly transitive subgroup of S_n, then $|G|$ is divisible by $n(n-1)$.

3.53. Let V be the plane, i.e., a real vector space of dimension 2. Show that $GL(2, \mathbf{R})$ is transitive on $V - \{0\}$. Let X be the semicircle consisting of all vectors $\alpha = e^{i\theta}$, where $0 \leq \theta < \pi$. Show that $GL(2, \mathbf{R})$ is doubly transitive on X.

Theorem 3.23 was used by G. Polya to solve some interesting combinatorial problems in chemistry; we illustrate his technique here. Given q distinct colors, how many striped neckties are there having n stripes (of equal width)? Clearly, the two neckties below are the same.



If, then, X is the set of all ordered sets of n colored boxes, then the cyclic group $G = \langle t \rangle$ operates on X, where t is the permutation

$$t = \begin{pmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{pmatrix},$$

and a necktie is just an orbit of G. In order to apply Theorem 3.33, we need only compute $F(1)$ and $F(t)$. There are q^n elements in X and each is fixed by the identity; hence $F(1) = q^n$. To compute $F(t)$, we first observe that, since t has order 2, Exercise 3.14 says that t is a product of disjoint transpositions. In fact, t is a product of k disjoint transpositions, where $k = [(n+1)/2]$, the greatest integer in $(n+1)/2$. If $n = 2k$, then $t = (1 \quad n)(2 \quad n-1)\cdots(k \quad k+1)$; if $n = 2k+1$, then $t = (1 \quad n)(2 \quad n-1)\cdots(k \quad k+2)$. An ordered set of $n = 2k$ colored boxes is thus fixed by t if and only if $q_1 = q_n$, $q_2 = q_{n-1}, \cdots, q_k = q_{k+1}$; there is a similar statement if $n = 2k+1$. We conclude that

$$F(t) = q^k = q^{[(n+1)/2]}.$$

The number of neckties is thus

$$\tfrac{1}{2}(q^n + q^{[(n+1)/2]}).$$

EXERCISES

3.54. If there are q colors available, prove that there are

$$\tfrac{1}{4}(q^{n^2} + 2q^{[(n^2+3)/4]} + q^{[(n^2+1)/2]})$$

distinct $n \times n$ colored tablecloths. (HINT: The group is a cyclic group of order 4 consisting of rotations of 0°, 90°, 180°, and 270°. The generator of the group is a permutation of n^2 ordered colored boxes and it is a product of disjoint 4-cycles.)

***3.55.** If there are q colors available, prove that there are

$$\frac{1}{n}\sum_{d|n} \varphi\left(\frac{n}{d}\right) q^d$$

colored roulette wheels having n compartments. In this formula, φ is the Euler φ-function, and the summation ranges over all divisors d of n. (HINT: The group is a cyclic group of order n operating by rotating by multiples of $(360/n)°$. Using Exercise 3.12 for $\alpha = (1\ 2 \cdots n)$, show that there are

$$\frac{1}{n}\sum_{0 \le k < n} q^{(n,k)}$$

roulette wheels. The desired formula arises from this one by collecting terms having the same exponent.)

We recommend the books of Burnside and Wielandt (see Bibliography) to the reader who wishes to study these ideas further.

CHAPTER 4

Finite Direct Products

DIRECT PRODUCTS

Definition If H and K are groups, the **(external) direct product** of H and K, denoted $H \times K$, is the set of all ordered pairs (h, k), where $h \in H$ and $k \in K$, with the binary operation

$$(h, k)(h', k') = (hh', kk').$$

It is easy to check that $H \times K$ is a group containing isomorphic copies of H and K, namely, $H \times \{1\}$ and $\{1\} \times K$.

EXERCISES

4.1. The operation of direct product is commutative and associative: for any groups H, K, and L,

$$H \times K \cong K \times H$$

and

$$(H \times K) \times L \cong H \times (K \times L).$$

Conclude that the notations $H_1 \times \cdots \times H_m$ and $\prod_{i=1}^{m} H_i$ are unambiguous.

4.2. $H \times \{1\}$ and $\{1\} \times K$ are normal subgroups of $H \times K$; these two subgroups generate $H \times K$, their intersection is $\{(1, 1)\}$, and the elements $(h, 1)$ and $(1, k)$ commute.

4.3. Let $\mu: G \times G \to G$ be the binary operation of a group G. If $G \times G$ is regarded as a group (the direct product), prove that μ is a homomorphism if and only if G is abelian.

4.4. Let $G = H \times K$ and let N be a normal subgroup of $H \times \{1\}$. Then N is a normal subgroup of G.

4.5. Prove that $Z(\prod_{i=1}^{m} H_i) = \prod_{i=1}^{m} Z(H_i)$. Conclude that a direct product of groups is abelian if and only if each factor is abelian.

***4.6.** Let G be an abelian group and let $f: H \to G$ and $g: K \to G$ be homomorphisms. Prove that there exists a unique homomorphism $F: H \times K \to G$ that extends both f and g, i.e., $F(h, 1) = f(h)$ and $F(1, k) = g(k)$. Show that this may be false if G is not abelian.

Notation If n is a positive integer, then $\sigma(n)$ will denote the cyclic group of order n.

Such a notation is legitimate, for we have already proven in Exercise 2.28 that any two cyclic groups of order n are isomorphic.

EXERCISES

****4.7.** If $(m, n) = 1$, then $\sigma(mn) \cong \sigma(m) \times \sigma(n)$.

4.8. If p is prime, prove that $\sigma(p^2) \ncong \sigma(p) \times \sigma(p)$.

We now adopt another point of view. It is easy to multiply two polynomials together; it is harder to factor a given polynomial. We have just seen how to "multiply" two groups together; can we "factor" a given group?

Theorem 4.1 *Let G be a group with normal subgroups H and K; if $H \cap K = \{1\}$ and $HK = G$, then $G \cong H \times K$.*

Proof Let $a \in G$. Since $HK = G$, $a = hk$ for some $h \in H$ and $k \in K$, and we claim that h and k are uniquely determined by a. If $a = h_1k_1$, then $hk = h_1k_1$ and $h^{-1}h_1 = kk_1^{-1}$. This element is simultaneously in H and K, i.e., in $H \cap K = \{1\}$; hence, $h = h_1$ and $k = k_1$.

Define $f: G \to H \times K$ by $f(a) = (h, k)$, where $a = hk$. Is f a homomorphism? If $a = hk$ and $a' = h'k'$, then $aa' = hkh'k'$, which is not in the proper form for evaluating f. Were it true that $kh' = h'k$, however, then we would be done. Indeed, we prove that for any $h \in H$ and $k \in K$, $hk = kh$. Consider the commutator $h^{-1}k^{-1}hk$. Now $(h^{-1}k^{-1}h)k \in K$, since K is normal, and $h^{-1}(k^{-1}hk) \in H$, since

H is normal. Therefore, this commutator is in $H \cap K = \{1\}$, and so $hk = kh$. We let the reader prove that f is a one-to-one correspondence. ∎

We pause to give an example that shows that all the hypotheses in Theorem 4.1 are necessary. Let $G = S_3$, $H = A_3$, and $K = \langle(12)\rangle$. It is easy to check that $H \cap K = \{1\}$ and $HK = G$. Note that H is normal, but that K is not. Were $S_3 = H \times K$, then $S_3 \cong \sigma(3) \times \sigma(2)$ which is abelian, a contradiction.

Theorem 4.2 *Let $G = H \times K$, and let $H_1 \lhd H$ and $K_1 \lhd K$. Then $H_1 \times K_1 \lhd G$ and*

$$G/(H_1 \times K_1) \cong (H/H_1) \times (K/K_1).$$

Proof Let $\pi: H \to H/H_1$ and $\rho: K \to K/K_1$ be the natural maps. Define $F: G \to (H/H_1) \times (K/K_1)$ by $F(h, k) = (\pi h, \rho k)$. The kernel of F is $H_1 \times K_1$ and the image of F is $(H/H_1) \times (K/K_1)$. ∎

Corollary 4.3 *If $G = H \times K$, then $G/(H \times \{1\}) \cong K$.*

The elements of an external direct product are ordered pairs, a rather restrictive condition. We say that a group G is the **(internal) direct product** of H and K if H and K are normal subgroups of G with $H \cap K = \{1\}$ and $HK = G$. The emphasis here is that the factors themselves, not merely isomorphic copies of them, lie in G. (If $G = H \times K$ is an external direct product, then it is also the internal direct product of $H \times \{1\}$ and $\{1\} \times K$, but it is not the internal direct product of H and K.) The two versions of direct product, of course, yield isomorphic groups. In the future, we shall not distinguish between external and internal, and we shall use "direct product" without an adjective. In almost all cases, however, our point of view is internal. For example, we shall write Corollary 4.3 as $(H \times K)/H \cong K$.

It is possible that $G = H \times K = H \times L$ (internal direct products) with $K \neq L$ (of course, $K \cong L$, by Corollary 4.3). For example, let $G = \langle a\rangle \times \langle b\rangle$, where $a^2 = 1 = b^2$. Then it is easy to see that G has a second decomposition: $G = \langle a\rangle \times \langle ab\rangle$. Now $\langle b\rangle \neq \langle ab\rangle$; in fact, $\langle b\rangle \cap \langle ab\rangle = \{1\}$. Therefore, the only uniqueness one has any right to expect of the factors in a direct product decomposition is up to isomorphism. We remark that there may be decompositions of a group in which the factors are not determined up to isomorphism, and, to end on a more optimistic note, there are some decompositions in which the factors are unique.

EXERCISES

*4.9. Denote by 0 the trivial homomorphism that sends every element into the identity. Prove that $G \cong H \times K$ if and only if there exist homomorphisms

$$H \underset{q}{\overset{i}{\rightleftarrows}} G \underset{j}{\overset{p}{\rightleftarrows}} K$$

with $qi = 1_H$ (the identity function on H), $pj = 1_K$, $pi = 0$, $qj = 0$, and

$$i(q(x))j(p(x)) = x \qquad \text{for all } x \in G.$$

4.10. Let G be a group with normal subgroups H and K. Prove that $G = H \times K$ (internal direct product) if and only if each $a \in G$ has a unique expression $a = hk$, where $h \in H$ and $k \in K$.

**4.11. If G is a group with normal subgroups $H_1, \cdots, H_m$, then $G = \prod_{i=1}^{m} H_i$ (internal) if and only if $G = \langle \bigcup_{i=1}^{m} H_i \rangle$ and, for all j, $H_j \cap \langle \bigcup_{i \neq j} H_i \rangle = \{1\}$.

4.12. Let $N \lhd G = H \times K$. Prove that either N is abelian or N intersects one of the factors H or K nontrivially.

4.13. Give an example of an abelian group $H \times K$ that contains a nontrivial subgroup N such that $N \cap H = \{1\}$ and $N \cap K = \{1\}$. Conclude that if $N \subset H \times K$, then $N \neq (N \cap H) \times (N \cap K)$ is possible.

THE BASIS THEOREM

For the next three sections, we shall deal exclusively with abelian groups. As is the usual custom, we now shift from multiplicative notation to additive notation. The following dictionary should prove useful.

ab	$a + b$
1	0
a^{-1}	$-a$
a^n	na
ab^{-1}	$a - b$
HK	$H + K$
aH	$a + H$
direct product	direct sum
$H \times K$	$H \oplus K$
$\prod_{i=1}^{m} H_i$	$\sum_{i=1}^{m} H_i$

If $G = H \times K$, H is called a **direct factor** of G; in additive notation, H is a **direct summand** of G.

Two remarks that hold for abelian groups greatly simplify our study:

1. If $a, b \in G$ and $n \in Z$, then $n(a + b) = na + nb$.
2. If X is a nonempty subset of G, then $\langle X \rangle$ is the set of all (finite) linear combinations of elements of X with coefficients in Z (Exercise 2.1).

Definition Let p be a prime. A group G is ***p*-primary** (or is a ***p*-group**) in case every element in G has order a power of p.

If one is working wholly in the context of abelian groups, he uses the term *p-primary*; otherwise, the usage of *p-group* is preferred.

Theorem 4.4 (Primary Decomposition) *Every finite abelian group G is a direct sum of p-primary groups.*

Proof For any prime p, let G_p be the set of all elements in G whose order is a power of p. Now $0 \in G_p$, and since G is abelian, G_p is a subgroup of G. We claim that $G = \sum G_p$, where the indices range over all primes p dividing $|G|$; we use the criterion of Exercise 4.11.

(i) Let $x \in G$, $x \neq 0$, and let the order of x be n. By the fundamental theorem of arithmetic, $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, where the p_i are distinct primes and the exponents $e_i \geq 1$. Set $n_i = n/p_i^{e_i}$, and observe that $(n_1, n_2, \cdots, n_k) = 1$ (for what prime could be a common divisor of the n_i?).[1] Therefore, there exist integers m_i such that $\sum m_i n_i = 1$; hence, $\sum (m_i n_i x) = x$. Note that $p_i^{e_i}(m_i n_i x) = m_i n x = 0$, so that $m_i n_i x \in G_{p_i}$. We conclude that the collection of G_p generates G.

(ii) Suppose $x \in G_p \cap \langle \bigcup_{q \neq p} G_q \rangle$. On the one hand, $p^e x = 0$ for some e; on the other hand, $x = \sum x_q$, where $q^{e_q} x_q = 0$ for exponents e_q. If we set $t = \prod q^{e_q}$, then $tx = 0$. Clearly, $(p^e, t) = 1$, so that there are integers a and b with $ap^e + bt = 1$. Therefore, $x = ap^e x + btx = 0$. ∎

Definition The subgroups G_p of G are called the **primary components** of G.

[1] A common divisor of a set of integers $\{n_1, \cdots, n_k\}$ is a nonzero integer c that divides each n_i; the greatest common divisor (gcd), denoted $(n_1, \cdots, n_k)$, is a positive common divisor that is divisible by every common divisor. It may be shown, in a manner analogous to the special case of a set of two integers, that the gcd exists and is a linear combination of $n_1, \cdots, n_k$ (see Appendix V).

We plan to show that every finite abelian group is a direct sum of cyclic groups. Theorem 4.4 allows us to consider, without loss of generality, the special case of finite p-primary abelian groups.

EXERCISE

**4.14. Let G be a p-primary abelian group and let $y_1, \cdots, y_t$ be nonzero elements of G for which

$$\langle y_1, \cdots, y_t \rangle = \langle y_1 \rangle \oplus \cdots \oplus \langle y_t \rangle.$$

(a) If $z_1, \cdots, z_t$ are elements of G with $pz_i = y_i$, for all i, then

$$\langle z_1, \cdots, z_t \rangle = \langle z_1 \rangle \oplus \cdots \oplus \langle z_t \rangle.$$

(b) If $k_1, \cdots, k_t$ are integers with $k_i y_i \neq 0$, then

$$\langle k_1 y_1, \cdots, k_t y_t \rangle = \langle k_1 y_1 \rangle \oplus \cdots \oplus \langle k_t y_t \rangle.$$

Definition Let G be an abelian group and m a positive integer.

$$mG = \{mx : x \in G\}.$$

Lemma 4.5 *An abelian group G with $pG = \{0\}$ is a vector space over Z_p, and it is a direct sum of cyclic groups of order p when G is finite.*

Proof Let $\bar{k}$ denote the congruence class of the integer k in Z_p. Define a scalar multiplication on G by

$$\bar{k}x = kx, \qquad \text{where } x \in G.$$

This operation is well defined, for if $k \equiv k' \bmod(p)$, then $k - k' = mp$ for some integer m, so that

$$(k - k')x = mpx = 0;$$

hence,

$$kx = k'x.$$

It is easily checked that G is a vector space over Z_p, and as such has a basis $\{x_1, x_2, \cdots, x_t\}$ when G is finite. We let the reader prove, using Exercise 4.11, that G is the direct sum of the $\langle x_i \rangle$. ∎

Theorem 4.6 (Basis Theorem) *Every finite abelian group G is a direct sum of primary cyclic groups.*

Proof By Theorem 4.4, we may assume that G is p-primary. We perform an induction on m, where m is an integer such that $p^m G = \{0\}$. If $m = 1$, the theorem is just Lemma 4.5.

Suppose that $p^{m+1}G = \{0\}$. If $H = pG$, then $p^mH = \{0\}$, so that induction gives

$$H = pG = \sum \langle y_i \rangle.$$

Since $y_i \in pG$, there are elements $z_i \in G$ with $pz_i = y_i$. If L is the subgroup of G generated by the z_i, then Exercise 4.14(a) shows that

$$L = \sum \langle z_i \rangle.$$

We interrupt the proof to understand the next step. Suppose the theorem were true, so that $G = \sum C_k$, where each C_k is cyclic. In considering pG, we are neglecting all C_k of order p (if any), for they were destroyed by multiplication by p. Now the construction of L has recaptured the C_k of order bigger than p, but nothing has yet been done to revive those C_k of order exactly p. Let us resume the proof.

We claim that L is a direct summand of G, and so we produce a subgroup M of G such that $L \oplus M = G$.

If $G[p] = \{x \in G: px = 0\}$, then $p(G[p]) = \{0\}$, so that $G[p]$ is a vector space over Z_p, by Lemma 4.5. If k_i is the order of y_i, then k_iz_i has order p, and so

$$k_iz_i \in G[p].$$

By Exercise 4.14(b), the set of k_iz_i is an independent subset of the vector space $G[p]$. Therefore, we can extend this set to a basis of $G[p]$, i.e, there are elements $\{x_1, x_2, \cdots, x_s\}$ such that

$$\{\text{the } k_iz_i, x_1, x_2, \cdots, x_s\}$$

is a basis of $G[p]$. Let $M = \langle x_1, x_2, \cdots, x_s \rangle$. Observe that, as in the proof of Lemma 4.5,

$$M = \sum \langle x_i \rangle.$$

(Of course, M is comprised of the resurrected summands of order p.)

We now show that $G = L \oplus M$, which will complete the proof.

(i) $L \cap M = \{0\}$. If $x \in L \cap M$, then $x = \sum b_iz_i = \sum a_jx_j$. Now $px = 0$ (since $x \in M$), so that $\sum pb_iz_i = 0$, which implies

$$0 = pb_iz_i = b_iy_i, \qquad \text{for all } i.$$

Hence, $b_i = b_i'k_i$ for some integers b_i', by Exercise 2.16. Therefore,

$$0 = \sum b_i'k_iz_i - \sum a_jx_j,$$

so that independence implies each term is 0; thus, $x = 0$.

(ii) $G = L + M$. Let $x \in G$. Now

$$px = \sum c_i y_i = \sum pc_i z_i,$$

so that

$$p(x - \sum c_i z_i) = 0,$$

and

$$x - \sum c_i z_i \in G[p].$$

Therefore,

$$x - \sum c_i z_i = \sum a_j x_j + \sum b_i k_i z_i,$$

i.e.,

$$x = \sum (c_i + b_i k_i) z_i + \sum a_j x_j \quad \in L + M. \ \blacksquare$$

Corollary 4.7 *Every finite abelian group G can be decomposed into a direct sum of cyclic groups*

$$G = \sigma(m_1) \oplus \sigma(m_2) \oplus \cdots \oplus \sigma(m_s),$$

where $m_i \mid m_{i+1}$ for all $i = 1, \cdots, s - 1$.

Proof G has a primary decomposition

$$G = G_{p_1} \oplus \cdots \oplus G_{p_t},$$

where $p_1, \cdots, p_t$ are distinct primes. By the basis theorem, we may assume each primary component is a direct sum of cyclic groups. For each i, let $\sigma(p_i^{e_i})$ be a cyclic summand in the given decomposition of G_{p_i} having largest order. Note that if $\sigma(p_i^{f_i})$ is any other cyclic summand in the decomposition, $f_i \leq e_i$, so that $p_i^{f_i} \mid p_i^{e_i}$. Since direct sum is commutative and associative, we have

$$G = H \oplus (\sigma(p_1^{e_1}) \oplus \cdots \oplus \sigma(p_t^{e_t})),$$

where H is the direct sum of the remaining cyclic summands. By Exercise 4.7,

$$G = H \oplus \sigma(m),$$

where $m = \prod p_i^{e_i}$. Repeating this process on H, we may write $H = H_1 \oplus \sigma(m_1)$. Moreover, $m_1 \mid m$ because we chose primary summands of largest order. Clearly this process ends in a finite number of steps. $\blacksquare$

Call a decomposition of G as in Corollary 4.7 a **canonical decomposition.**

EXERCISES

4.15. Let G be a finite abelian group of order n. If m divides n, show that G contains a subgroup of order m. (Compare with Theorem 3.11.)

***4.16.** A finite p-primary abelian group is generated by its elements of largest order.

***4.17.** Let F be a finite field and let G be the multiplicative group consisting of its nonzero elements. Prove that G is cyclic. (HINT: Suppose G is not cyclic; use Theorem 4.6 and the theorem stating that a polynomial of degree k over F has at most k roots in F.) Generalize to any finite multiplicative subgroup of the group of nonzero elements of a field.

****4.18.** If $G = \sum_{i=1}^{n} H_i$, then $mG = \sum_{i=1}^{n} mH_i$.

4.19. If $G = \sum_{i=1}^{n} H_i$, then $G[p] = \sum_{i=1}^{n} (H_i[p])$.

4.20. Let G be a finite p-primary abelian group and let $x \in G$ have order maximal among the elements of G. Prove that $\langle x \rangle$ is a direct summand of G.

****4.21.** Let $G = \sigma(m_1) \oplus \cdots \oplus \sigma(m_s)$ be a canonical decomposition. Show that $|G| = \prod m_i$ and that m_s is the least postive integer n for which $nG = \{0\}$. (The least $n > 0$ with $nG = \{0\}$ is called the **exponent** of G; this terminology is also used for nonabelian groups.)

***4.22.** Let H be a finite abelian group with $pH = \{0\}$ for some prime p (H is called an **elementary** abelian group). Prove that any two decompositions of H into a direct sum of cyclic groups have the same number of summands. Denote this number $d(H)$. (The astute reader will note that $d(H)$ is the dimension of H considered as a vector space over Z_p.)

****4.23.** Let G be a direct sum of b copies of cyclic groups of order p^k. If $n < k$, then $d(p^nG/p^{n+1}G) = b$.

****4.24.** If H and K are elementary p-primary abelian groups, then $d(H \oplus K) = d(H) + d(K)$.

THE FUNDAMENTAL THEOREM OF FINITE ABELIAN GROUPS

We now have quite a bit of information about finite abelian groups, but we still have not answered the basic question: If G and H are finite abelian groups, when are they isomorphic? Since both G and H are direct

sums of cyclic groups, your first guess is that $G \cong H$ if they have the same number of summands of each kind. There are two things wrong with this guess. First of all, since, e.g., $\sigma(6) \cong \sigma(3) \oplus \sigma(2)$, we had better require that G and H have the same number of *primary* summands of each kind. Our second objection is much more serious. How can we count summands at all? To do so would require a unique factorization theorem analogous to the fundamental theorem of arithmetic, where the analog of a prime number is a primary cyclic group. Such an analog does exist; it is called the fundamental theorem of finite abelian groups, and it is this theorem we now discuss.

We have already seen, in Exercise 4.22, that the number of cyclic summands occurring in a decomposition of an elementary abelian group depends only on G and not on the particular decomposition; this number is denoted $d(G)$. For any nonnegative integer n and any finite p-primary abelian group G, the quotient group $p^nG/p^{n+1}G$ is elementary, so that $d(p^nG/p^{n+1}G)$ is defined.

Lemma 4.8 *Let $G = \sum C_i$ be a decomposition of a finite p-primary abelian group into a direct sum of cyclic groups. Then $d(p^nG/p^{n+1}G)$ is the number of cyclic summands C_i having order $\geq p^{n+1}$.*

Proof Let B_k denote the direct sum of all those C_i in the decomposition having order exactly p^k, and let b_k be the number of C_i in B_k (of course, b_k may be zero). Thus

$$G = B_1 \oplus \cdots \oplus B_t.$$

By Exercise 4.18,

$$p^nG = p^nB_{n+1} \oplus p^nB_{n+2} \quad \oplus \cdots \oplus p^nB_t$$

and

$$p^{n+1}G = \qquad\qquad p^{n+1}B_{n+2} \oplus \cdots \oplus p^{n+1}B_t.$$

By Theorem 4.2,

$$p^nG/p^{n+1}G \cong p^nB_{n+1} \oplus (p^nB_{n+2}/p^{n+1}B_{n+2}) \oplus \cdots \oplus p^nB_t/p^{n+1}B_t.$$

Hence $d(p^nG/p^{n+1}G) = b_{n+1} + b_{n+2} + \cdots + b_t$, by Exercises 4.23 and 4.24. ∎

Definition If G is a finite p-primary abelian group, then

$$U(n, G) = d(p^nG/p^{n+1}G) - d(p^{n+1}G/p^{n+2}G).$$

For each n, the integer $U(n, G)$ depends only on G and not on any particular choice of decomposition of G into a direct sum of cyclic groups.

Theorem 4.9 *Let G be a finite p-primary abelian group. Any two decompositions of G into a direct sum of cyclic groups have the same number of summands of each order. In fact, the number of cyclic summands of order p^{n+1} is $U(n, G)$.*

Proof Choose a decomposition of G. It follows from Lemma 4.8 that $U(n, G)$ is the number of cyclic summands of order exactly p^{n+1}. The first statement now follows, because $U(n, G)$ does not depend on the choice of decomposition. ∎

Corollary 4.10 *Let G and H be finite p-primary abelian groups. Then $G \cong H$ if and only if $U(n, G) = U(n, H)$ for all $n \geq 0$.*

Proof It is easy to check that an isomorphism between G and H induces isomorphisms $p^nG/p^{n+1}G \cong p^nH/p^{n+1}H$ for all n. Hence $U(n, G) = U(n, H)$ for all n.

Conversely, $G \cong H$ because they have direct sum decompositions into cyclic groups of the same type, by Theorem 4.9. ∎

We have only to delete the adjective "p-primary" in Theorems 4.9 and 4.10 to finish our discussion. The hard work has already been done, and the following three theorems (whose proofs may be supplied by the reader) complete the picture.

Lemma 4.11 *Let G and H be finite abelian groups, and let $f: G \to H$ be a homomorphism. For each p,*

$$f(G_p) \subset H_p.$$

Theorem 4.12 *Let G and H be finite abelian groups; $G \cong H$ if and only if $G_p \cong H_p$ for all primes p.*

Theorem 4.13 (Fundamental Theorem of Finite Abelian Groups) *Let G be a finite abelian group. Any two decompositions of G into direct sums of primary cyclic groups have the same number of summands of each order.*

Corollary 4.14 *Let G be a finite abelian group. If*

$$G = \sigma(m_1) \oplus \cdots \oplus \sigma(m_s)$$

and

$$G = \sigma(n_1) \oplus \cdots \oplus \sigma(n_t)$$

are canonical decompositions, then $s = t$ and $m_i = n_i$ for all i.

Proof By Exercise 4.21, $m_s = n_t$, for both equal the exponent of G. The proof is completed by induction on $\max\{s, t\}$, for the complementary summands are isomorphic, by the fundamental theorem. ∎

EXERCISES

4.25. Let G and H be finite p-primary abelian groups. Then, for all $n \geq 0$,

$$U(n, G \oplus H) = U(n, G) + U(n, H).$$

4.26. Let G and H be finite abelian groups. If $G \oplus G \cong H \oplus H$, prove that $G \cong H$.

4.27. Suppose A, B, and C are finite abelian groups. If $A \oplus B \cong A \oplus C$, then $B \cong C$.

4.28. How many nonisomorphic abelian groups are there of order 360?

***4.29.** Let H be a subgroup of a finite abelian group G. Prove that G contains a subgroup isomorphic to G/H.

4.30. Let G and H be finite abelian groups such that, for each k, both G and H have the same number of elements of order k. Then $G \cong H$.

MODULES AND MATRICES

We digress from our study of groups to apply Theorems 4.6 and 4.13 to linear algebra; we shall prove the existence and uniqueness of the rational canonical form of a square matrix over an arbitrary field F. At this stage, this project is one of translation, so that we need only introduce a new vocabulary. Our exposition is complete, but since we are assuming that the reader is comfortable with linear algebra, our pace is not leisurely.

Definition Let R be a commutative ring with unit. An **ideal** I in R is a nonempty subset of R such that

(i) $a, b \in I$ imply $a - b \in$ I.
(ii) $a \in I$ and $r \in R$ imply that $ra \in I$.

An important example of an ideal is the set of all multiples of a fixed $r_0 \in R$ by elements of R; this ideal is denoted (r_0) and is called the **principal ideal generated** by r_0. Thus,

$$(r_0) = \{x \in R: x = rr_0 \quad \text{for some } r \in R\}.$$

Definition A **principal ideal domain** is a domain in which every ideal is a principal ideal (a **domain** is a commutative ring with unit that has no divisors of zero).

Example 1 In Z, condition (ii) follows from (i); since every subgroup of Z is cyclic, every ideal in Z is principal.

Example 2 If F is any field, then the only ideals in F are $\{0\}$ and F itself; since $\{0\} = (0)$ and $F = (1)$, F is a principal ideal domain.

Example 3 Let F be a field and let $F[x]$ be the ring of polynomials in x with coefficients in F. The reader may prove that any nonzero ideal I in $F[x]$ consists precisely of all the multiples (by polynomials) of the monic polynomial (i.e., leading coefficient is 1) of least degree which is in I.

Many properties of the ring of integers can be generalized to any principal ideal domain R. An element $u \in R$ is a **unit** if there is an element $v \in R$ with $uv = 1$. In Z, the only units are ± 1; in F, every nonzero element is a unit; in $F[x]$, the nonzero constants are the units. A nonzero element $p \in R$ is **irreducible** if p is not a unit, and in every factorization $p = ab$ in R, either a or b is a unit. In Z, the irreducibles are the primes (positive and negative); in F, there are no irreducibles; in $F[x]$, the irreducibles are the irreducible polynomials.

A **common divisor** of $a_1, \cdots, a_n \in R$ is an element $c \in R$ that divides each a_i (i.e., $cb_i = a_i$ for some $b_i \in R$); a **greatest common divisor**, gcd, of $a_1, \cdots, a_n$ is a common divisor that is divisible by every common divisor.

The following two theorems are proved in Appendix V.

Theorem A *Every finite set of elements $a_1, \cdots, a_n$ in a principal ideal domain R has a gcd; moreover, a gcd is a linear combination of $a_1, \cdots, a_n$ with coefficients in R.*

Theorem B *If d and d' are gcd's of $a_1, \cdots, a_n$ in a principal ideal domain R, then d and d' are* **associates**, *that is, $d' = ud$ for some unit $u \in R$.*

Because of Theorem B, we may choose gcd's in Z to be positive and gcd's in $F[x]$ to be monic polynomials.

The fundamental theorem of arithmetic may be generalized to principal ideal domains (a proof is in Appendix V):

Theorem F *If R is a principal ideal domain, then*

(i) *every nonzero $a \in R$ that is not a unit is a product of irreducible elements;*

(ii) *this factorization is unique in the following sense: if $p_1 \cdots p_m = q_1 \cdots q_n$, where the p and q are irreducibles, then there is a one-to-one correspondence between the factors ($m = n$) with corresponding factors associate.*

We shall need a generalization of Z_p, the integers modulo p.

Theorem 4.15 (Construction of a Quotient Ring) *Let R be a commutative ring with unit and let I be an ideal in R. There exists a ring S and a ring homomorphism π of R onto S whose kernel is I.*

Proof Under addition, R is an abelian group and I is a subgroup. Therefore, R/I is an additive abelian group, and the natural map $\pi: R \to R/I$ is a group homomorphism of R onto R/I whose kernel is I.[2]

In order that $S = R/I$ be a ring, we define

$$(r + I)(r' + I) = rr' + I.$$

This is a well-defined multiplication, for if $r + I = s + I$ and $r' + I = s' + I$, then $rr' - ss' = r(r' - s') + (r - s)s' \in I$; hence, $rr' + I = ss' + I$. The reader should verify that R/I is a ring under the given operations. Finally, since $\pi(r) = r + I$, π is a ring homomorphism. ∎

Definition The ring just constructed is denoted R/I and is called the **quotient ring** of R modulo I.

EXERCISES

4.31. If $R = Z$ and $I = (m)$, then the ring R/I is isomorphic to the ring Z_m.

4.32. The first isomorphism theorem (Theorem 2.12) holds for commutative rings: If $f: R \to S$ in a ring homomorphism with kernel I, then I is an ideal and $R/I \cong$ image f.

****4.33.** The correspondence theorem (Theorem 2.16) holds for commutative rings, if one replaces "normal subgroup" by "ideal".

[2] The zero element in R/I is thus $I = 0 + I$.

**4.34. If R is a commutative ring with unit, then R is a field if and only if R contains no proper ideals.

Definition An ideal I in R is **prime** in case $I \neq R$ and $rr' \in I$ implies either r or r' lies in I.

EXERCISE

**4.35. If R is a principal ideal domain, then the nonzero prime ideals are the ideals (p), where p is irreducible.

Theorem 4.16 *Let R be a commutative ring with unit. An ideal I in R is a prime ideal if and only if R/I is a domain.*

Proof If I is a prime ideal, we must show that R/I contains no zero-divisors. Suppose $(r + I)(r' + I) = 0$, i.e., $rr' + I = I$. Then $rr' \in I$; since I is a prime ideal, one of these factors, say r, lies in I. Hence, $r + I = 0$.

Suppose R/I is a domain. If $rr' \in I$, then $(r + I)(r' + I) = 0$ in R/I, so that one of the factors is 0, i.e., either r or r' lies in I. Thus I is a prime ideal. ∎

Definition Let R be a commutative ring with unit. An ideal I in R is a **maximal ideal** in case $I \neq R$ and there is no larger proper ideal of R that contains I.

Theorem 4.17 *Let R be a commutative ring with unit. An ideal I in R is a maximal ideal if and only if R/I is a field.*

Proof If I is a maximal ideal, then the correspondence theorem for rings (Exercise 4.33) implies that R/I has no proper ideals. By Exercise 4.34, R/I is a field. To prove the converse, just reverse this argument. ∎

Corollary 4.18 *Every maximal ideal in R is a prime ideal.*

In general, the converse of this corollary is false. For example, if $R = Z[x]$, one may verify that (x) is a prime ideal that is contained in the proper ideal consisting of all polynomials in $Z[x]$ having even constant term. However, the converse of the corollary is true when R is a principal ideal domain.

Theorem 4.19 *If R is a principal ideal domain, every nonzero prime ideal I is a maximal ideal.*

Proof Let J be an ideal with $I \subsetneq J$. Since R is a principal ideal domain, there are elements a and b in R with $I = (a)$ and $J = (b)$. Now $a \in J = (b)$, so there is an $r \in R$ with $a = rb$. Since I is a prime ideal, either $r \in I$ or $b \in I$. Were b in I, then $J \subset I$, a contradiction. Therefore, $r \in I$, so that $r = sa$ for some $s \in R$. Hence, $a = rb = sab$, and $1 = sb$. The ideal (b) thus contains 1, and so $(b) = R$. We conclude that I is a maximal ideal. ∎

Corollary 4.20 *If R is a principal ideal domain and $p \in R$ is irreducible, then $R/(p)$ is a field.*

Proof (p) is a prime ideal, by Exercise 4.35, which is maximal, by Theorem 4.19. ∎

Definition An abelian group V is an ***R*-module** in case a "scalar multiplication" is defined, i.e., there is a function $R \times V \to V$ (whose values we write in multiplicative notation) that satisfies:

$$\left.\begin{array}{ll} \text{(i)} & (rs)\alpha = r(s\alpha); \\ \text{(ii)} & (r + s)\alpha = r\alpha + s\alpha; \\ \text{(iii)} & r(\alpha + \beta) = r\alpha + r\beta; \\ \text{(iv)} & 1\alpha = \alpha \end{array}\right\} \text{for every } \alpha, \beta \in V \text{ and } r, s, 1 \in R.$$

Thus, an R-module is just like a vector space except that the scalars are allowed to be in a ring R instead of a field.

Example 4 If $R = Z$, an R-module is an abelian group, for axioms (i) to (iv) are always true for scalars in Z.

Example 5 If $R = F$, an R-module is a vector space over F.

Example 6 Let V be a vector space over F, and let $T: V \to V$ be a linear transformation. We make V into an $F[x]$-module, denoted V^T, by defining

$$(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)\alpha = a_0\alpha + a_1T\alpha + a_2T^2\alpha + \cdots + a_nT^n\alpha$$

(T^n is the composite of T with itself n times). The reader should check that we have defined a scalar multiplication.

Just as a principal ideal domain R is a generalization of Z, so are R-modules generalizations of abelian groups. Almost any theorem that

can be proved for abelian groups has a true analog for R-modules; moreover, the proofs of the theorems for R-modules are exact translations of the proofs for abelian groups. Here is the dictionary from the language of groups to that of modules.

A **submodule** of an R-module V is a subgroup W of V that is closed under scalar multiplication: if $\alpha \in W$ and $r \in R$, then $r\alpha \in W$.

Let W be a submodule of V. Remember that modules are just abelian groups with added structure. If we forget the added structure for a moment, then W is a subgroup of the abelian group V, and so V/W is an abelian group. Define the **quotient module** V/W to be the abelian group V/W with scalar multiplication

$$r(v + W) = rv + W.$$

EXERCISES

4.36. R itself is an R-module. Prove that the submodules of R are its ideals.

****4.37.** Let V be a vector space over F, let $T: V \to V$ be a linear transformation, and let V^T be the $F[x]$-module of Example 6. Prove that $W \subset V$ is a submodule if and only if W is a subspace of V with $T(W) \subset W$. (W is called an **invariant subspace**.)

4.38. If V is an R-module and $\alpha \in V$, then

$$\langle \alpha \rangle = \{r\alpha : r \in R\}$$

is a submodule of V.

4.39. The intersection of any family of submodules of V is again a submodule of V.

4.40. If X is a subset of an R-module V, then there is a smallest submodule W of V containing X, which we denote $\langle X \rangle$. Further, if X is nonempty,

$$\langle X \rangle = \{\sum r_i x_i : r_i \in R \text{ and } x_i \in X\}.$$

An R-module V is **finitely generated** in case it contains a finite number of elements $\alpha_1, \alpha_2, \cdots, \alpha_k$ such that every element in V is a linear combination of these α with coefficients in R. In particular, an R-module V is **cyclic** if it can be generated by one of its elements.

If V and W are R-modules, their **direct sum** $V \oplus W$ is the usual direct sum made into an R-module by $r(v, w) = (rv, rw)$. There is also an internal version: If V is an R-module with submodules W_1 and W_2, then V is the **(internal) direct sum** if $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$.

EXERCISES

4.41. If V is an R-module with submodules $W_1, \cdots, W_n$, then $V = \sum W_i$ (internal) if and only if $V = W_1 + \cdots + W_n$ and, for all j, $W_j \cap \langle \bigcup_{i \neq j} W_i \rangle = \{0\}$.

****4.42.** An abelian group G is finite if and only if G is finitely generated and every element in G has finite order.

Exercise 4.42 tells us how to translate "finite abelian group" once we can translate "order" of an element. Before doing this, we note that Exercise 4.42 is false if we do not assume that the group G is abelian. Burnside proved that if a finitely generated group G comprised of elements of finite order can be imbedded in $GL(n, \mathbf{C})$, where $\mathbf{C}$ is the complex numbers, then G is finite. In 1964, Golod and Šafarevič exhibited an infinite finitely generated p-group; in 1970, Adjan and Novikov, together, and Britton, independently, exhibited an infinite finitely generated group of finite exponent (so this latter group has a uniform bound on the orders of its elements).

Let $\alpha \in V$, where V is an R-module. The **order ideal** of $\alpha = \{r \in R: r\alpha = 0\}$. It is quickly verified that the order ideal is an ideal of R. Because R is a principal ideal domain, this order ideal consists of all the multiples of a fixed element in R. In Z, we choose this fixed element to be positive (and we get the usual definition of order); in $F[x]$, we choose this fixed element to be the monic polynomial of smallest degree in the ideal. Thus, if $\alpha \in V^T$, its order ideal is generated by the monic polynomial $m(x)$ of least degree for which $m(T)\alpha = 0$.

We say that $\alpha \in V$ has **finite order** if its order ideal is nonzero. Finally, an R-module is ***p*-primary** if the order ideal of each element is generated by some power of an irreducible element p.

Theorem 4.21 *Let R be a principal ideal domain and let V be a finitely generated R-module in which every element has finite order. Then*

$$V = \langle \alpha_1 \rangle \oplus \cdots \oplus \langle \alpha_s \rangle.$$

Moreover, the cyclic summands may be chosen to satisfy either of the following conditions. If the order ideal of α_i is (r_i),

(i) *each r_i is a power of an irreducible element in R;*

or

(ii) $r_i \mid r_{i+1}$ for $i = 1, \cdots, s - 1$.

Proof Our proofs of the corresponding group theorems were written so that translation into module language is mechanical. Thus, the cyclic decomposition as in (i) is Theorems 4.4, 4.5, and 4.6, while the cyclic decomposition as in (ii) is Corollary 4.7. ▮

Call a decomposition of an R-module V as in Theorem 4.21(ii) a **canonical decomposition** of V.

Theorem 4.22 *Let V be a finite dimensional vector space over F and let $T\colon V \to V$ be a linear transformation. Then*

$$V = W_1 \oplus \cdots \oplus W_s,$$

where each $W_i = \langle \alpha_i \rangle$ is a cyclic invariant subspace. Moreover, the cyclic invariant spaces may be chosen to satisfy either of the following conditions. If the order ideal of α_i is $(f_i(x))$,

(i) *each $f_i(x)$ is a power of an irreducible polynomial;*

or

(ii) $f_i(x) \mid f_{i+1}(x)$ for $i = 1, \cdots, s - 1$.

Proof Recall that we may regard V as an $F[x]$-module V^T. Since V is finite dimensional, it has a basis $\{\beta_1, \cdots, \beta_n\}$, so that each vector $\gamma \in V$ may be written

$$\gamma = \sum c_i \beta_i, \qquad c_i \in F.$$

It follows that $\{\beta_1, \cdots, \beta_n\}$ generates V^T as an $F[x]$-module, for now we are allowed to use as coefficients any polynomial over F, not merely constant ones (of course, we may not now need all the β_i to generate V^T). Further, every $\gamma \in V$ is annihilated by some polynomial (this follows, for example, from the Cayley-Hamilton theorem); therefore, every γ has finite order.

We have verified that the hypotheses of Theorem 4.21 hold, so V^T may be decomposed as a direct sum of cyclic submodules. Since a submodule of V^T is just an invariant subspace (Exercise 4.37), the theorem is true. ▮

EXERCISES

**4.43. A subspace W of V^T is a cyclic invariant subspace if and only if there is a vector $\alpha \in W$ such that $\{\alpha, T\alpha, T^2\alpha, \cdots, T^{k-1}\alpha\}$ is a basis of W for some $k \geq 1$.

4.44. If W has a basis of the form

$$\{\alpha, T\alpha, T^2\alpha, \cdots, T^{k-1}\alpha\},$$

then adjoining $T^k\alpha$ to this set makes it dependent, and there is an equation

$$T^k\alpha = \sum_{i=0}^{k-1} b_i T^i\alpha.$$

Prove that the order ideal of α is generated by the polynomial

$$x^k - b_{k-1}x^{k-1} - \cdots - b_0.$$

We remind the reader of the correspondence between linear transformations on V and matrices. Let V be m-dimensional and let $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m\}$ be an ordered basis of V. If T: $V \to V$, then, for each i, $T\varepsilon_i$ is a linear combination of the ε_j:

$$T\varepsilon_i = \sum a_{ji}\varepsilon_j.$$

The matrix of T relative to the original basis of the ε_j is $A = (a_{ji})$. Therefore, the coordinates of $T\varepsilon_1$ form the first *column* of A, the coordinates of $T\varepsilon_2$ form the second *column* of A, and so forth.

Definition Let $c(x) = x^k - b_{k-1}x^{k-1} - \cdots - b_0$. The **companion matrix** of $c(x)$ is the $k \times k$ matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & b_0 \\ 1 & 0 & 0 & \cdots & 0 & b_1 \\ 0 & 1 & 0 & \cdots & 0 & b_2 \\ 0 & 0 & 1 & \cdots & 0 & b_3 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_{k-1} \end{bmatrix}.$$

EXERCISES

****4.45.** Let T: $V \to V$ and let W be a cyclic invariant subspace of V. If $\{\alpha, T\alpha, T^2\alpha, \cdots, T^{k-1}\alpha\}$ is a basis of W, the matrix of T on W is the $k \times k$ companion matrix of $c(x)$, where $c(x)$ generates the order ideal of α.

4.46. Let A be the companion matrix of $c(x)$. Prove that the characteristic polynomial of A is $c(x)$.

Definition Let A be a $k \times k$ matrix and let B be an $m \times m$ matrix. The **direct sum** of A and B is the $(k + m) \times (k + m)$ matrix

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

The analogous definition for the direct sum of a finite number of matrices may be supplied by the reader. Note that the direct sum of A and B is similar to the direct sum of B and A.

Theorem 4.23 *Every $n \times n$ matrix A over a field F is similar to a direct sum of companion matrices $C_1, \cdots, C_s$; if $f_i(x)$ is the characteristic polynomial of C_i, we may assume either*

(i) *each $f_i(x)$ is a power of an irreducible polynomial;*

or

(ii) $f_i(x) \mid f_{i+1}(x)$, $i = 1, \cdots, s - 1$.

Proof Let V be an n-dimensional vector space over F; choose some ordered basis of V. The matrix A now defines a linear transformation $T: V \to V$, for its columns tell us how to transform the chosen basis of V. Given this T, we make V into an $F[x]$-module V^T. Now V is a direct sum of cyclic invariant subspaces W_i, by Theorem 4.22. If α_i is a generator of W_i, then Exercise 4.43 says that a new basis of V is

$$\{\alpha_1, T\alpha_1, T^2\alpha_1, \cdots; \alpha_2, T\alpha_2, T^2\alpha_2, \cdots; \cdots; \alpha_s, T\alpha_s, T^2\alpha_s, \cdots\}.$$

The matrix B of T with respect to this new basis is a direct sum of companion matrices, by Exercise 4.45. Finally, A and B are similar, for they represent the same linear transformation relative to different bases of V. ∎

A matrix B that is a direct sum of companion matrices as in Theorem 4.23(ii) is called a **rational canonical form**; the polynomials of the companion matrices are called the **elementary divisors** of B. More generally, if a matrix A is similar to a rational canonical form B, then the elementary divisors of A are defined to be the elementary divisors of B. (The question whether A can be similar to different rational canonical forms will be answered shortly.)

Recall that the **minimum polynomial** of an $n \times n$ matrix A is the monic polynomial $m(x)$ of least degree for which $m(A) = 0$.

EXERCISES

4.47. Let A have elementary divisors $f_1(x), \cdots, f_s(x)$, where $f_i(x) \mid f_{i+1}(x)$, $i = 1, \cdots, s-1$. Prove that the characteristic polynomial of A is $\prod f_i(x)$ and that the minimum polynomial of A is $f_s(x)$. (Comparison with Exercise 4.21 shows that the characteristic polynomial is thus analogous to the order of a group whereas the minimum polynomial is analogous to the exponent.)

4.48. If A is a companion matrix, its minimum polynomial is its characteristic polynomial.

In Chapter 8, we shall study certain groups whose elements are nonsingular matrices. Since the order of an element is the same as the order of any of its conjugates, we can compute the order of a matrix by computing the order of its canonical form. Unfortunately, it is difficult to compute powers of companion matrices. If the field F is a large one, however, there is another canonical form whose powers are easily calculated.

Definition A $k \times k$ **Jordan block** is a $k \times k$ matrix of the form

$$\begin{bmatrix} a & 0 & 0 & \cdots & 0 & 0 \\ 1 & a & 0 & \cdots & 0 & 0 \\ 0 & 1 & a & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & a & 0 \\ 0 & 0 & 0 & \cdots & 1 & a \end{bmatrix}.$$

(A 1×1 Jordan block is a 1×1 matrix (a).)

EXERCISES

****4.49.** Prove that for every positive integer k,

$$\begin{bmatrix} b & 0 \\ 1 & b \end{bmatrix}^k = \begin{bmatrix} b^k & 0 \\ kb^{k-1} & b^k \end{bmatrix}.$$

****4.50.** Prove that for every positive integer k,

$$\begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}^k = \begin{bmatrix} a^k & 0 & 0 \\ ka^{k-1} & a^k & 0 \\ s(k)a^{k-2} & ka^{k-1} & a^k \end{bmatrix}$$

where $s(k) = k(k-1)/2$.

**4.51. Let $T: W \to W$ and let $\{\alpha, T\alpha, \cdots, T^{k-1}\alpha\}$ be a basis of W, so that the matrix of T is the companion matrix of some polynomial $f(x)$. Suppose $f(x) = (x - a)^k$. If

$$\beta_0 = \alpha, \beta_1 = (T - aE)\alpha, \cdots, \beta_{k-1} = (T - aE)^{k-1}\alpha,$$

prove that $\{\beta_0, \beta_1, \cdots, \beta_{k-1}\}$ is a basis of W (E is the identity transformation).

Theorem 4.24 *Let A be an $n \times n$ matrix over F, where F is a field that contains all the characteristic roots of A. Then A is similar to a direct sum of Jordan blocks.*

Proof After looking at the primary decomposition in Theorem 4.23(i), we see that it suffices to prove that a companion matrix C whose characteristic polynomial $f(x)$ is a power of an irreducible polynomial is similar to a Jordan block.

Our hypothesis on F tells us that $f(x) = (x - a)^k$ for some $a \in F$. Now C determines a transformation $T: W \to W$, where W is a vector space with a basis of the form $\{\alpha, T\alpha, \cdots, T^{k-1}\alpha\}$. Let us compute the matrix of T relative to the basis $\{\beta_0, \beta_1, \cdots, \beta_{k-1}\}$ of W that we examined in Exercise 4.51.

If $j + 1 < k$,

$$\begin{aligned} T\beta_k &= T(T - aE)^j\alpha \\ &= (T - aE)^j T\alpha \\ &= (T - aE)^j[(T - aE) + aE]\alpha \\ &= (T - aE)^{j+1}\alpha + (T - aE)^j a\alpha \\ &= \beta_{j+1} + a\beta_j. \end{aligned}$$

If $j + 1 = k$, then $(T - aE)^{j+1} = (T - aE)^k = 0$, by the Cayley-Hamilton theorem. Therefore, $T\beta_{k-1} = a\beta_{k-1}$.

The matrix of T is thus a Jordan block J, and so C and J are similar. ∎

A direct sum of Jordan blocks is called a **Jordan canonical form**. Thus Theorem 4.24 says that if F contains enough elements, a square matrix is similar to a Jordan canonical form. For example, if F is the complex numbers (or any algebraically closed field), then every square matrix is similar to a Jordan canonical form.

The uniqueness of the various canonical forms will follow from a translation of the fundamental theorem of finite abelian groups.

Definition If V and W are R-modules, a function $f: V \to W$ is an **R-homomorphism** in case

$$f(v + v') = f(v) + f(v')$$

and

$$f(rv) = rf(v)$$

for all $v, v' \in V$ and all $r \in R$. An R-homomorphism $f: V \to W$ is an **R-isomorphism** if it is a one-one correspondence; in this case, we say V and W are **R-isomorphic** and write $V \cong W$.

Example 7 If $R = Z$, then V and W are merely abelian groups. Every homomorphism $f: V \to W$ is a Z-homomorphism, for $f(mv) = mf(v)$ for every integer m.

Example 8 If R is a field, V and W are vector spaces and R-homomorphisms are linear transformations.

Example 9 If $R = F[x]$, then an R-homomorphism $f: V \to W$ has the property that

$$f(\psi(x)v) = \psi(x)f(v)$$

for every polynomial $\psi(x) \in F[x]$.

EXERCISES

4.52. Prove the first isomorphism theorem for modules (quotient modules have been defined earlier). (HINT: Since modules are groups, the reader need only check that the group isomorphism in Theorem 2.12 is, in fact, an R-isomorphism.)

4.53. Every cyclic R-module V is isomorphic to R/I, where I is the order ideal of a generator of V. Conclude that two cyclic modules are isomorphic if and only if they have generators with the same order ideal.

4.54. Let R be a principal ideal domain and let $a, b \in R$ be relatively prime (gcd = 1). Then

$$R/(a) \oplus R/(b) \cong R/(ab).$$

In words, a direct sum of cyclic modules of relatively prime orders is cyclic.

Theorem 4.25 (Fundamental Theorem for Modules) *Let R be a principal ideal domain and let V be a finitely generated R-module in which every element has finite order.*

(i) *Any two decompositions of V into direct sums of primary cyclic modules have the same number of summands of each order.*

(ii) *Any two canonical decompositions of V have the same number of cyclic summands of each order.*

Proof Translation of Theorem 4.13 and Corollary 4.14. ∎

In order to apply this fundamental theorem to matrices, we continue our examination of $F[x]$-modules.

Lemma 4.26 *Let V and W be vector spaces over F and let $T: V \to V$ and $S: W \to W$ be linear transformations. A function $f: V^T \to W^S$ is an $F[x]$-homomorphism if and only if:*

(i) *f is a linear transformation of the vector spaces V and W;*

(ii) *$f(T(v)) = S(f(v))$ for all $v \in V$.*

Proof If f is an $F[x]$-homomorphism, then $f(\psi(x)v) = \psi(x)f(v)$ for all $v \in V$ and all polynomials $\psi(x) \in F[x]$. In particular, if $\psi(x)$ is a constant polynomial, then we see that f is a linear transformation; if $\psi(x) = x$, then $f(xv) = xf(v)$. But the definition of scalar multiplication in V^T is $xv = T(v)$ and the definition of scalar multiplication in W^S is $xf(v) = S(f(v))$.

For the converse, we are told that

$$f(\psi(x)v) = \psi(x)f(v)$$

for all the polynomials $\psi(x) =$ constant and for the polynomial $\psi(x) = x$. It follows easily that the equation holds for all polynomials $\psi(x)$. ∎

Theorem 4.27 *If A and B are $n \times n$ matrices over a field F, then A is similar to B if and only if the corresponding $F[x]$-modules they determine are $F[x]$-isomorphic.*

Proof Let V be a vector space over F with basis $\{\alpha_1, \cdots, \alpha_n\}$ and let T and S be the linear transformations on V defined by A and B, respectively. We let V^T denote V made into an $F[x]$-module by $x\alpha = T\alpha$ and we let V^S denote V made into an $F[x]$-module by $x\alpha = S\alpha$.

If A is similar to B, there is a nonsingular matrix P with $PAP^{-1} = B$, and P defines a linear transformation $f: V \to V$. We claim that f is even an $F[x]$-isomorphism between V^T and V^S. By Lemma 4.26, it suffices to prove that $f(T(v)) = S(f(v))$ for all $v \in V$,

i.e., $fT = Sf$. In terms of matrices, this is the given equation $PA = BP$.

Suppose, conversely, that $f\colon V^T \to V^S$ is an $F[x]$-isomorphism. By Lemma 4.26, $Sf = fT$, or, since f is an isomorphism, $S = fTf^{-1}$. If P is the matrix corresponding to the linear transformation $f\colon V \to V$, then $B = PAP^{-1}$, i.e., A and B are similar. ∎

Theorem 4.28 *Two $n \times n$ matrices A and B over a field F are similar if and only if they have the same elementary divisors. An $n \times n$ matrix is similar to exactly one rational canonical form.*

Proof Only necessity needs proof. By the preceding theorem, A and B are similar if and only if the modules they determine are $F[x]$-isomorphic. It follows from part (ii) of the fundamental theorem for modules, Theorem 4.25, that the order ideals of cyclic summands in canonical decompositions are the same. But the monic generators $f_i(x)$ of these ideals are by definition the elementary divisors.

If A is similar to two rational canonical forms B_1 and B_2, then B_1 and B_2 are similar, hence have the same elementary divisors, hence are comprised of the same companion matrices. Therefore, $B_1 = B_2$. ∎

Theorem 4.29 *Let an $n \times n$ matrix A be similar to B_1 and B_2, each a direct sum of companion matrices whose polynomials are powers of irreducibles. Then B_1 and B_2 are the same except for a permutation of the companion matrices.*

For example, we may have

$$B_1 = \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \quad \text{and} \quad B_2 = \begin{bmatrix} C_2 & 0 \\ 0 & C_1 \end{bmatrix},$$

where C_1, C_2 are companion matrices.

Proof Exactly as the proof of Theorem 4.28 except that we use part (i) of the fundamental theorem for modules instead of part (ii). ∎

EXERCISES

4.55. Prove that a Jordan canonical form associated with an $n \times n$ matrix A is unique within a permutation of its Jordan blocks.

4.56. If b and b' are nonzero elements of a field F, then $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ and $\begin{bmatrix} a & b' \\ 0 & c \end{bmatrix}$ are similar.

4.57. Let A and B be $n \times n$ matrices with entries in a field F. Supposing that F is a subfield of a field K and that A and B are similar over K, prove that A and B are similar over F.

THE REMAK-KRULL-SCHMIDT THEOREM

If a nonabelian group G is a direct product of normal subgroups, each of which cannot be factored further, is this factorization unique? The affirmative answer for a large class of groups (which contains all finite groups) is the main result of this section. Since we shall consider nonabelian groups, we return to the multiplicative notation.

Definition A group G is **indecomposable** if $G \neq \{1\}$ and if $G \cong H \times K$, then either H or $K = \{1\}$.

EXERCISES

4.58. Which of the following groups are indecomposable? (a) Z; (b) the additive group of rationals; (c) S_n; (d) $\sigma(21)$; (e) the multiplicative group of positive rationals; (f) $\sigma(p^n)$, where p is prime.

4.59. Give necessary and sufficient conditions that a finite abelian group be indecomposable.

Definition A homomorphism $f: G \to G$ is called an **endomorphism** of G; an isomorphism $f: G \to G$ is called an **automorphism** of G.

There are certain endomorphisms of G that arise quite naturally in the consideration of direct products, as we saw in Exercise 4.9.

Definition Let $G = H_1 \times \cdots \times H_m$. The homomorphisms $\varepsilon_i: G \to H_i$ defined by $\varepsilon_i(h_1 h_2 \cdots h_m) = h_i$ are called **projections**.

Throughout this section, direct products are internal. Thus, each H_i is a normal subgroup of G and there is an inclusion map $\lambda_i: H_i \to G$. The composite $\lambda_i \varepsilon_i$ is then an endomorphism of G.

EXERCISES

4.60. Let ε_i be a projection and λ_i the inclusion (as in above paragraph). The endomorphism $\lambda_i \varepsilon_i$ is **idempotent**, i.e., $\lambda_i \varepsilon_i \circ \lambda_i \varepsilon_i = \lambda_i \varepsilon_i$.

4.61. An endomorphism β of G is **normal** in case

$$a^{-1}\beta(x)a = \beta(a^{-1}xa)$$

for every a and x in G. Prove that $\lambda_i\varepsilon_i$ is normal, where ε_i is a projection, λ_i the corresponding inclusion.

****4.62.** The composite of normal endomorphisms is a normal endomorphism.

****4.63.** If $\alpha: G \to G$ is a normal endomorphism and $H \lhd G$, then $\alpha(H) \lhd G$.

We now introduce a new way of combining endomorphisms; unfortunately, the new function we get is not always a homomorphism.

Definition If α and β are endomorphisms of G, then $\alpha + \beta: G \to G$ is the function defined by

$$(\alpha + \beta)(x) = \alpha(x)\beta(x) \qquad \text{for every } x \in G.$$

EXERCISES

4.64. If G is abelian and α and β are endomorphisms of G, then $\alpha + \beta$ is also an endomorphism of G.

4.65. Let α and β be endomorphisms of S_3 defined as follows: α is conjugation by (123) and β is conjugation by (132). Show that $\alpha + \beta$ is not an endomorphism of S_3.

****4.66.** If $G = H_1 \times H_2 \times \cdots \times H_n$ has projections $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ and inclusions $\lambda_i: H_i \to G$, then the sum of any k distinct $\lambda_i\varepsilon_i$ is a normal endomorphism of G. Moreover,

$$\lambda_1\varepsilon_1 + \cdots + \lambda_n\varepsilon_n = 1_G.$$

We now consider a condition on a group G that will ensure that it is a direct product of indecomposable groups (for there do exist groups without this property).

Definition A group G has the ACC **(ascending chain condition)** in case every increasing chain of normal subgroups of G stops; i.e., if

$$A_1 \subset A_2 \subset \cdots \subset A_n \subset A_{n+1} \subset \cdots$$

is a chain of normal subgroups of G, then there is an integer t for which $A_t = A_{t+1} = A_{t+2} = \cdots$.

A group G has the DCC **(descending chain condition)** in case every decreasing chain of normal subgroups of G

$$B_1 \supset B_2 \supset \cdots \supset B_n \supset B_{n+1} \supset \cdots$$

stops.

G **has both chain conditions** if it has both chain conditions!

Every finite group has both chain conditions; Z has the ACC but not the DCC; in Chapter 9, we shall study a group $\sigma(p^\infty)$ that has the DCC but not the ACC.

Lemma 4.30 *If G has either chain condition, then G is a direct product of a finite number of indecomposable groups.*

Proof Let us call a group *good* if it satisfies the conclusion of Lemma 4.30; otherwise it is *bad*. Clearly, an indecomposable group is good. If A and B are good, so is $A \times B$. Thus, if a group is bad, it is not indecomposable and it has a proper, bad factor.

Suppose now that G is bad. Set $G = A_0$. By induction, we may assume there exist bad subgroups $A_0, \cdots, A_n$ such that A_i is a direct factor of A_{i-1} for $i = 1, \cdots, n$. Since A_n is bad, it has a proper, bad factor A_{n+1}. By induction, we obtain a strictly decreasing infinite chain of normal subgroups of G:

$$G = A_0 > A_1 > A_2 > \cdots.$$

If G has the DCC, we have reached a contradiction.

Now assume G has the ACC. Since each A_i is a direct factor of A_{i-1}, we have normal subgroups B_i with $A_{i-1} = A_i \times B_i$. There is thus an increasing chain of normal subgroups

$$B_1 < B_1 \times B_2 < B_1 \times B_2 \times B_3 < \cdots$$

and we have reached a contradiction in this case, too. ∎

There do exist indecomposable groups having neither chain condition, e.g., the additive group of rationals.

Lemma 4.31 *Let G have both chain conditions and let α be a normal endomorphism of G. Then α is one-to-one if and only if α is onto. (Thus, either property ensures that α is an automorphism.)*

Proof Suppose α is one-to-one and $g \notin \alpha(G)$. We prove, by induction, that $\alpha^n(g) \notin \alpha^{n+1}(G)$. Otherwise there would be an element

$h \in G$ with $\alpha^n(g) = \alpha^{n+1}(h)$, so that $\alpha(\alpha^{n-1}(g)) = \alpha(\alpha^n(h))$. Since α is one-to-one, $\alpha^{n-1}(g) = \alpha^n(h) \in \alpha^n(G)$, which contradicts the inductive hypothesis. Thus, we have a strictly descending chain of subgroups

$$G \supset \alpha(G) \supset \alpha^2(G) \supset \cdots.$$

Since α is normal, each $\alpha^n(G)$ is a normal subgroup of G, and so the DCC is violated. Therefore, α is onto.

Suppose α is onto. Let $A_n = \text{kernel } \alpha^n$; each A_n is a normal subgroup of G because α^n is a homomorphism (the normality of α is here irrelevant). Thus, we have the ascending chain of normal subgroups:

$$\{1\} = A_0 \subset A_1 \subset A_2 \subset \cdots.$$

Since G satisfies the ACC, this chain stops. Let t be the smallest integer for which $A_t = A_{t+1} = \cdots$. We claim that $t = 0$, which will prove our theorem. If $t \geq 1$, there is an $x \in A_t$ with $x \notin A_{t-1}$, i.e., $a^t(x) = 1$, but $\alpha^{t-1}(x) \neq 1$. Since α is onto, there is an element $g \in G$ with $\alpha(g) = x$. Hence, $1 = \alpha^t(x) = \alpha^{t+1}(g)$, so that $g \in A_{t+1} = A_t$. Therefore, $\alpha^t(g) = 1$; but $\alpha^t(g) = \alpha^{t-1}(\alpha(g)) = \alpha^{t-1}(x)$, so that $\alpha^{t-1}(x) = 1$, which is a contradiction. Thus α is one-to-one. ∎

Definition An endomorphism α of G is **nilpotent** if there is a positive integer k such that $\alpha^k = 0$, where 0 denotes the trivial endomorphism that sends every element of G into the identity.

Theorem 4.32 (Fitting's Lemma) *Let G have both chain conditions and let α be a normal endomorphism of G. Then $G = K \times H$, where H and K are each invariant under α, $\alpha|K$ is nilpotent, and $\alpha|H$ is onto.*

Proof Let $K_n = \text{kernel } \alpha^n$ and $H_n = \text{image } \alpha^n$. As we observed above, there are two chains of normal subgroups:

$$G \supset H_1 \supset H_2 \supset \cdots \quad \text{and} \quad \{1\} \subset K_1 \subset K_2 \subset \cdots.$$

Since G has both chain conditions, each of these chains stops: the H_n after r steps and the K_n after s steps. Let t be the larger of r and s, so that $K_t = K_{t+1} = \cdots$ and $H_t = H_{t+1} = \cdots$; define $H = H_t$ and $K = K_t$. It is easy to check that H and K are each invariant under α.

Suppose $x \in H \cap K$. Now $x \in H$ implies $\alpha^t(g) = x$ for some $g \in G$, and $x \in K$ implies $\alpha^t(x) = 1$. Therefore, $\alpha^{2t}(g) = \alpha^t(x) = 1$, so that $g \in K_{2t} = K_t$. Hence, $\alpha^t(g) = 1$ and $x = 1$.

Let $g \in G$. Then $\alpha^t(g) \in H_t = H_{2t}$, so there is an $x \in G$ with $\alpha^{2t}(x) = \alpha^t(g)$. Applying α^t to $g\alpha^t(x^{-1})$ gives 1, so that $g\alpha^t(x^{-1}) \in K_t = K$. Hence, $g = [g\alpha^t(x^{-1})]\alpha^t(x) \in KH$, so that $G = K \times H$.

Now $\alpha(H) = \alpha(H_t) = \alpha(\alpha^t(G)) = \alpha^{t+1}(G) = H_{t+1} = H_t = H$, so that $\alpha|H$ is onto. Finally, let $x \in K$; then $\alpha^t(x) \in K \cap H = \{1\}$ so that $\alpha|K$ is nilpotent. ▮

Corollary 4.33 *Let G be an indecomposable group having both chain conditions. Any normal endomorphism α of G is either nilpotent or an automorphism.*

Proof By Theorem 4.32, $G = K \times H$ with $\alpha|K$ nilpotent and $\alpha|H$ onto. Since G is indecomposable, either $G = K$ or $G = H$. In the first case, α is nilpotent; in the second case, α is onto and hence is an automorphism, by Lemma 4.31. ▮

Theorem 4.34 *Let G be an indecomposable group having both chain conditions and let α_1 and α_2 be normal, nilpotent endomorphisms of G. If $\alpha_1 + \alpha_2$ is also an endomorphism of G, then it is nilpotent.*

Proof If $\alpha_1 + \alpha_2$ is an endomorphism of G, it is immediately seen to be normal, so that, by Corollary 4.33, it is either nilpotent or an automorphism. Suppose $\alpha_1 + \alpha_2$ is an automorphism and γ its inverse; γ is easily seen to be normal. Set $\lambda_1 = \alpha_1\gamma$ and $\lambda_2 = \alpha_2\gamma$, so that $1_G = \lambda_1 + \lambda_2$, i.e., $\lambda_1(x)\lambda_2(x) = x$ for all $x \in G$. In particular, $\lambda_1(x^{-1})\lambda_2(x^{-1}) = x^{-1}$. If we take the inverse of both sides, we see that $\lambda_2(x)\lambda_1(x) = x$, and so $\lambda_1 + \lambda_2 = \lambda_2 + \lambda_1$. Now the equation $\lambda_1(\lambda_1 + \lambda_2) = (\lambda_1 + \lambda_2)\lambda_1$ implies that $\lambda_1\lambda_2 = \lambda_2\lambda_1$. It follows that the set of all endomorphisms of G obtained from λ_1 and λ_2 by sums and products forms a commutative ring with unit. Hence, the binomial theorem applies: for any integer $m > 0$,

$$(\lambda_1 + \lambda_2)^m = \lambda_1^m + \binom{m}{1}\lambda_1^{m-1}\lambda_2 + \binom{m}{2}\lambda_1^{m-2}\lambda_2^2 + \cdots + \lambda_2^m.$$

Since α_1 and α_2 are nilpotent, λ_1 and λ_2 are nilpotent (they cannot be automorphisms since they have nontrivial kernels). Therefore, there are positive integers r and s with $\lambda_1^r = 0$ and $\lambda_2^s = 0$. If we take m large enough ($m = r + s - 1$ will do), then we obtain $(\lambda_1 + \lambda_2)^m = 0$. Since $\lambda_1 + \lambda_2 = 1_G$, we have $1_G = 0$, which contradicts the fact that, as any indecomposable group, $G \neq \{1\}$, ▮

Corollary 4.35 *Let G be an indecomposable group having both chain conditions; let $\alpha_1, \cdots, \alpha_n$ be a set of normal, nilpotent endomorphisms such that every sum of distinct α_i is an endomorphism. Then $\alpha_1 + \cdots + \alpha_n$ is nilpotent.*

Proof Induction on n. ∎

Theorem 4.36 (Remak-Krull-Schmidt)[3] *Let G be a group having both chain conditions. If*

$$G = H_1 \times H_2 \times \cdots \times H_s$$

and

$$G = K_1 \times K_2 \times \cdots \times K_t$$

are two decompositions of G into indecomposable groups, then $s = t$ and there is a reindexing so that $H_i \cong K_i$ for all i. Moreover, given any r between 1 and s, the reindexing may be chosen so that

$$G = H_1 \times \cdots \times H_r \times K_{r+1} \times \cdots \times K_t.$$

REMARK Our conclusion is stronger than saying that the factors appearing in the two factorizations are determined up to isomorphism; we can even replace factors in one decomposition with factors of the other.

Proof We shall give the proof for the case $r = 1$, leaving the rest of the proof for the reader to finish by induction. Given the first decomposition, we must find a renumbering of the K so that $H_i \cong K_i$ for all i and

$$G = H_1 \times K_2 \times \cdots \times K_t.$$

Let $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_s$ be the projections on the H and let $\eta_1, \eta_2, \cdots, \eta_t$ be the projections on the K. Let $\lambda_i\colon H_i \to G$ and $\mu_j\colon K_j \to G$ be inclusion maps. Then the maps $\lambda_i\varepsilon_i$ and $\mu_j\eta_j$ are endomorphisms of G.

Now Exercise 4.66 says that every partial sum of $\sum \mu_j\eta_j$ is a normal endomorphism of G. Hence any partial sum of

$$\sum \varepsilon_1\mu_j\eta_j\lambda_1 = \varepsilon_1 \circ \sum \mu_j\eta_j \circ \lambda_1$$

is a normal endomorphism of H_1. Corollary 4.35 implies that not every $\varepsilon_1\mu_j\eta_j\lambda_1$ is nilpotent. By Corollary 4.33, one of these maps must be an automorphism of H_1. We renumber so that $\varepsilon_1\mu_1\eta_1\lambda_1$ is an automorphism; let γ be its inverse.

[3] This theorem was first given by Wedderburn, and it has since been polished and generalized by Remak, Krull, Schmidt, Fitting, Ore, Kuroš, Azumaya, Jónsson, Atiyah, and others.

We shall now show that the map $\eta_1\lambda_1: H_1 \to K_1$ is an isomorphism. We do know that the following composite ψ is the identity:

$$H_1 \xrightarrow{\gamma} H_1 \xrightarrow{\lambda_1} G \xrightarrow{\eta_1} K_1 \xrightarrow{\mu_1} G \xrightarrow{\varepsilon_1} H_1.$$

Consider the normal composite $\theta: K_1 \to K_1$:

$$K_1 \xrightarrow{\mu_1} G \xrightarrow{\varepsilon_1} H_1 \xrightarrow{\gamma} H_1 \xrightarrow{\lambda_1} G \xrightarrow{\eta_1} K_1.$$

Using the fact that the composite $\psi: H_1 \to H_1$ above is the identity, it follows easily that θ is idempotent: $\theta \circ \theta = \theta$. Since K_1 is indecomposable with both chain conditions, θ is either nilpotent or an automorphism, by Corollary 4.33. Thus, either $\theta = 0$ or $\theta = 1_{K_1}$. But $\theta \neq 0$ because θ occurs in the composite $\psi \circ \psi = \psi = 1_{H_1}$. Therefore θ is the identity and $\eta_1\lambda_1: H_1 \to K_1$ is an isomorphism (with inverse $\gamma\varepsilon_1\mu_1$). It follows that $\varepsilon_1\mu_1: K_1 \to H_1$ is an isomorphism, for γ is an isomorphism.

Now η_1 sends $K_2 \times \cdots \times K_t$ into 1, but $\eta_1\lambda_1$ is an isomorphism on H_1. Therefore,

$$H_1 \cap (K_2 \times \cdots \times K_t) = \{1\}.$$

If we define $G^* = \langle H_1, K_2 \times \cdots \times K_t \rangle$, then

$$G^* = H_1 \times K_2 \times \cdots \times K_t.$$

If $x \in G$, then $x = k_1k_2 \cdots k_t$, where each $k_j \in K_j$. Since $\varepsilon_1\mu_1$ is an isomorphism, the map $\beta: G \to G$ defined by $\beta(x) = \varepsilon_1(k_1)k_2 \cdots k_t$ is one-to-one and has image G^*. By Lemma 4.31, β must be onto and so $G^* = G$. Finally,

$$K_2 \times \cdots \times K_t \cong G/H_1 \cong H_2 \times \cdots \times H_s,$$

so that the remainder of the theorem follows by induction on $\max\{s, t\}$. ∎

EXERCISES

4.67. Use the Remak-Krull-Schmidt theorem to prove the fundamental theorem of finite abelian groups.

4.68. If $K \lhd G$ and if both K and G/K have both chain conditions, prove that G has both chain conditions. Conclude that if K and H each have both chain conditions, so does $K \times H$. (HINT: Use the Dedekind law, Exercise 2.55).

4.69. Let G have both chain conditions; if there is a group H with $G \times G \cong H \times H$, then $G \cong H$.

**4.70. Let G have both chain conditions. If $G \cong A \times B$ and $G \cong A \times C$, then $B \cong C$.

4.71. Let G be the additive group of $Z[x]$. Prove that $G \times Z \cong G \times Z \times Z$. Conclude that the "cancellation law" is not always valid.

The Remak-Krull-Schmidt theorem holds for algebraic systems other than groups with both chain conditions; in particular, it holds for R-modules having both chain conditions (where the phrase "normal subgroup" is replaced by "submodule" in the definitions of ACC and DCC). The proof of the Remak-Krull-Schmidt theorem given above is also a proof for R-modules if one further replaces "homomorphism" by "R-homomorphism", and "normal endomorphism" by "R-endomorphism". As an illustration of the value of this translation, we state Fitting's lemma for vector spaces (i.e., R-modules, where R is a field).

Fitting's Lemma for Vector Spaces *If V is a finite-dimensional vector space and $T: V \to V$ is a linear transformation, then $V = W_1 \oplus W_2$, where W_1 and W_2 are invariant under T, $T|W_1$ is nonsingular, and $T|W_2$ is nilpotent.*

The matrix version of Fitting's lemma thus says: Every $n \times n$ matrix over a field F is similar to a matrix of the form

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix},$$

where A is nilpotent and B is nonsingular.

The proofs of generalizations of the Remak-Krull-Schmidt theorem to more general situations, e.g., to groups with operators or to lattices, are easily accessible in the literature; cf. A. G. Kurosh, *The Theory of Groups*, Vol. II (see Bibliography).

CHAPTER

5

The Sylow Theorems

p-GROUPS

We recall the following definition from the preceding chapter.

Definition If p is a prime, a group G is a ***p*-group** in case every element $x \in G$ has order some power of p.

We begin with a characterization of finite p-groups (Corollary 5.3).

Lemma 5.1 *If G is a finite abelian group whose order is divisible by a prime p, then G contains an element of order p.*

Proof[1] Let $x \in G$ be distinct from 1. If the order of x is pm, then the reader can easily check that x^m is an element in G of order p.

Suppose that x has order t, where $(p,t) = 1$. Since G is abelian, $\langle x \rangle$ is a normal subgroup and $G/\langle x \rangle$ is an abelian group of order $|G|/t$. Since $|G|/t$ is divisible by p and is less than $|G|$, an induction on $|G|$ provides an element $\bar{y} \in G/\langle x \rangle$ of order p. Hence, if $y \in G$ goes into $\bar{y}$ under the natural map, the order of y is a multiple of p (Exercise 2.29). We have returned to the first case. ∎

Theorem 5.2 (Cauchy) *If G is a finite group whose order is divisible by a prime p, then G contains an element of order p.*

[1] This lemma follows immediately from the basis theorem (Theorem 4.6), but we give an argument from first principles so that its elementary character is not disguised.

Proof If $x \in G$, the number of conjugates of x is $[G: C(x)]$, where $C(x)$ is the centralizer of x in G. Now if $x \notin Z(G)$, then $|C(x)| < |G|$, so that if p divides $|C(x)|$, we are done, by induction. Therefore, we may assume that p does not divide $|C(x)|$ for all $x \notin Z(G)$; better, since $|G| = [G: C(x)]|C(x)|$, we may assume that p does divide $[G: C(x)]$ for all $x \notin Z(G)$ (it is here we use the fact that p is prime).

We partition G into its conjugacy classes and count:

$$(*) \qquad |G| = |Z(G)| + \sum [G: C(x)],$$

where the summation ranges over a complete set of nonconjugate x not in $Z(G)$. Now p divides both $|G|$ and $\sum [G: C(x)]$, so that p divides $|Z(G)|$. Since $Z(G)$ is an abelian group, it contains an element of order p, by Lemma 5.1. ∎

The equation (*) is called the **class equation**.

Corollary 5.3 *A finite group G is a p-group if and only if $|G|$ is a power of p.*

Proof Suppose there is a prime $q \neq p$ dividing $|G|$. By Cauchy's theorem, there is an element $a \in G$ of order q, contradicting the fact that G is a p-group.

The converse follows immediately from Lagrange's theorem (Theorem 2.8). ∎

Theorem 5.4 *If G is a finite p-group with more than one element, then the center of G, $Z(G)$, has more than one element.*

Proof Partitioning G into its conjugacy classes gives the class equation:

$$|G| = |Z(G)| + h_1 + \cdots + h_m,$$

where the h_i are the sizes of the distinct conjugacy classes of noncentral elements of G. Suppose $a \in G$ is not in $Z(G)$; then $C(a)$ is a proper subgroup of G, and so, by Corollary 5.3, $[G: C(a)]$ is a power of p. Thus, p divides each h_i and, therefore, also $|Z(G)|$. ∎

Corollary 5.5 *Any group G of order p^2 (p a prime) is abelian.*

Proof Suppose G is not abelian, i.e., $G \neq Z(G)$. Then $Z(G)$ has order p (for this order cannot be 1). Therefore, $G/Z(G)$ has order p and so is cyclic, contradicting Exercise 3.22. ∎

EXERCISES

**5.1. Let $H \lhd G$. If both H and G/H are p-groups, then G is a p-group.

**5.2. Let $|G| = p^n$, where p is prime. If $0 \leq k \leq n$, prove that G contains a normal subgroup of order p^k.

5.3. Let G be a finite p-group and let $\{1\} \neq H \lhd G$. Prove that $H \cap Z(G) \neq \{1\}$.

5.4. Let G be a finite p-group and let H be a proper subgroup of order p^k. Then H can be imbedded in a subgroup of order p^{k+1}. Conclude that every maximal subgroup of G has index p.

THE SYLOW THEOREMS

The major results in this section are basic for understanding the structure of a finite group. We shall prove the existence of largest possible p-subgroups of a finite group G. Now several such subgroups may exist, and we can count the number of them (within a congruence). On the other hand, these subgroups are unique in the sense that any two are isomorphic via a conjugation of G.

Definition Let p be a prime. A subgroup P of G is a **p-sylow[2] subgroup** of G if it is a maximal p-subgroup of G.

Thus, if P is a p-sylow subgroup of G and if Q is a p-subgroup of G such that $Q \supset P$, then $Q = P$. Observe that every p-subgroup of G is contained in a p-sylow subgroup of G. (If G is infinite, this statement may be proved using Zorn's lemma, Appendix IV.)

We have seen the utility of the notion of conjugate and centralizer of an element; we now consider the analogs of these notions for a subgroup.

Definition Let H be a subgroup of G. A subgroup S of G is a **conjugate** of H in case there is an element $a \in G$ with $S = aHa^{-1}$. One often writes $S = H^a$.

Observe that conjugate subgroups are isomorphic.

Definition Let H be a subgroup of G. The **normalizer** of H in G, denoted $N_G(H)$, is the set

$$N_G(H) = \{a \in G : aHa^{-1} = H\}.$$

[2] After L. Sylow (1832–1918).

When the meaning is clear from the context, we shall abbreviate $N_G(H)$ by $N(H)$. It is easy to see that $N_G(H)$ is a subgroup of G containing H, and it is the largest such subgroup in which H is normal.

EXERCISES

5.5. If either H is finite or a has finite order, then $a \in N_G(H)$ if and only if $aHa^{-1} \subset H$. (There exist infinite groups in which aHa^{-1} is a proper subgroup of H.)

5.6. If X is the set of all subgroups of G, then G is a permutation group on X acting as conjugations. If $H \in X$, then the orbit of H is the set of all subgroups conjugate to H, and the stabilizer of H is $N_G(H)$.

5.7. The number of distinct conjugates of H in G is $[G: N(H)]$, and so this number divides $|G|$ if G is finite.

5.8. Let G be a finite group with proper subgroup H. Prove that G is not the set-theoretical union of all the conjugates of H. Give an example in which H is not normal and this union is a subgroup.

****5.9.** Every conjugate of a p-sylow subgroup of G is itself a p-sylow subgroup of G. Conclude that if, for some fixed prime p, G has only one p-sylow subgroup P, then P is normal in G.

5.10. Find a 2-sylow subgroup of S_4. (HINT: Consider the symmetries of a square.)

Lemma 5.6 *Let P be a p-sylow subgroup of G. Then $N(P)/P$ has no elements $\neq 1$ whose order is a power of p.*

Proof Suppose $\bar{x} \in N(P)/P$ has order a power of p. If S^* is the subgroup of $N(P)/P$ generated by $\bar{x}$, then S^* is a p-group. By the correspondence theorem (Theorem 2.16), there is a subgroup S of G containing P such that $S/P \cong S^*$. By Exercise 5.1, S is a p-group containing P, so that the maximality of P implies $S = P$. Therefore, $S^* = \{1\}$ and $\bar{x} = 1$. ∎

Lemma 5.7 *Let P be a p-sylow subgroup of G and let $a \in G$ have order a power of p. If $aPa^{-1} = P$, then $a \in P$.*

Proof Clearly, $a \in N(P)$. If $\pi: N(P) \to N(P)/P$ is the natural map, then $\pi(a)$ has order a power of p (since a has). By Lemma 5.6, $\pi(a) = 1$, i.e., $a \in \ker \pi = P$. ∎

We recall the following fact about orbits.

Lemma 5.8 *Let P be a p-group and let $\alpha: P \to S_n$ be a homomorphism. The size of each orbit of $\alpha(P)$ is a power of p (which may be $p^0 = 1$).*

Proof Since P is a p-group, $\alpha(P)$ is a finite p-group and so has order a power of p. By Corollary 3.22, the size of each orbit divides $|\alpha(P)|$. ∎

We are ready to prove the main theorems of this chapter.

Theorem 5.9 (Sylow) *Let G be a finite group with p-sylow subgroup P. All p-sylow subgroups of G are conjugate to P, and the number of these subgroups is a divisor of $|G|$ that is $\equiv 1$ (mod p).*

Proof The basic idea is the realization (Exercise 5.9) that conjugation by any element of G sends a p-sylow subgroup into a p-sylow subgroup.

Let X be the set of all the conjugates of P: say,

$$X = \{P_1, P_2, \cdots, P_r\}$$

(where our notation sets $P = P_1$). For any $a \in G$, we define $\alpha_a: X \to X$ by

$$\alpha_a(P_i) = aP_ia^{-1}.$$

Now each α_a is one-to-one, for if $aP_ia^{-1} = aP_ja^{-1}$, then $P_i = P_j$. Since any function on a finite set that is one-to-one must be onto, each α_a is a permutation of X. Even more is true: The function $\alpha: G \to S_X \cong S_r$ given by $a \to \alpha_a$ is a homomorphism:

$$\alpha_a\alpha_b(P_i) = \alpha_a(bP_ib^{-1}) = abP_ib^{-1}a^{-1} = (ab)P_i(ab)^{-1} = \alpha_{ab}(P_i),$$

so that $\alpha_a\alpha_b = \alpha_{ab}$.

Restrict the map α to a homomorphism $P \to S_X$. By Lemma 5.8, the size of each orbit of $\alpha(P)$ is a power of p. Now what does it mean to say that one of these sizes is 1? There would be an i with $\alpha_a(P_i) = P_i$ for all α_a, i.e, $aP_ia^{-1} = P_i$ for all $a \in P$. By Lemma 5.7, for all $a \in P = P_1$, $a \in P_i$, so that $P \subset P_i$. Since P is a p-sylow subgroup, we must have $P = P_i$. Conclusion: Each orbit has size an "honest" power of p save $\{P_1\}$, which has size 1. Therefore, $r \equiv 1 \pmod{p}$.

Suppose now that Q is a p-sylow subgroup of G which is not a conjugate of P, i.e., Q is not a P_i for any i. Restrict the map α to a homomorphism of $Q \to S_X$. Again Lemma 5.8 tells us that the size of each orbit of $\alpha(Q)$ is a power of p, and again we ask if any of these have size 1. The same argument as before shows that if there were such an orbit $\{P_i\}$, then $Q = P_i$, and this is contrary to the choice of

Q. But now every orbit has size an honest power of p, so that p divides r, i.e., $r \equiv 0 \pmod{p}$. This contradicts our previous congruence, so that no such Q can exist. Therefore, every p-sylow subgroup of G is conjugate to P.

Finally, the number of conjugates of P is the index of its normalizer, and so it is a divisor of $|G|$. ∎

It follows that the converse of Exercise 5.9 is true: If a group G has a normal p-sylow subgroup P, then P is the unique p-sylow subgroup of G.

Theorem 5.10 (Sylow) *Let G be a finite group of order $p^k m$, where $(p, m) = 1$. Every p-sylow subgroup P of G had order p^k.*

Proof We first prove that $[G: P]$ is prime to p. Since $[G: P] = [G: N(P)][N(P): P]$, it suffices to prove that each of these factors is prime to p. Now $[G: N(P)]$ is just the number of conjugates of P, which we have just seen is $\equiv 1 \pmod{p}$. Also, $[N(P): P] = |N(P)/P|$, and $N(P)/P$ has no elements of order p; by Cauchy's theorem, $|N(P)/P|$ is prime to p. It follows that $[G: P]$ is prime to p.

By Lagrange's theorem, $|P| = p^n$ for $n \leq k$, so that $|G|/|P| = mp^{k-n}$. But $|G|/|P| = [G: P]$ is prime to p, so that $k = n$. ∎

Corollary 5.11 *Let G be a finite group and let p be a prime. If p^n divides $|G|$, then G contains a subgroup of order p^n.*

Proof If P is a p-sylow subgroup of G, then p^n divides $|P|$, by the preceding theorem. By Exercise 5.2, P (hence G) contains a subgroup of order p^n. ∎

We have now seen how much of the converse of Lagrange's theorem can be salvaged. If m divides $|G|$ and m is a power of a prime, then G contains a subgroup of order m; if m has two distinct prime factors, however, we have exhibited a group G, namely A_4, such that m divides $|G|$ and such that G contains no subgroup of order m.

EXERCISES

5.11. (Frattini Argument) Let H be a subgroup of K, where $K \lhd G$. If, for every $g \in G$, there exists $k \in K$ with $gHg^{-1} = kHk^{-1}$, then

$$G = KN_G(H).$$

5.12. Let $K \lhd G$, where G is finite, and let P be a p-sylow subgroup of K. Then

$$G = KN_G(P).$$

(This result is often called the Frattini argument too.)

****5.13.** Let G be a finite group with a p-sylow subgroup P. Prove that any subgroup of G that contains $N_G(P)$ (inclusion not necessarily proper) is equal to its own normalizer.

****5.14.** Let G be a finite group, and suppose that, for every prime p, every p-sylow subgroup of G is normal. Prove that G is the direct product of its sylow subgroups. (Compare with Theorem 4.4.)

5.15. Let G be a finite group with normal subgroup H, and let P be a p-sylow subgroup of G. Prove that $H \cap P$ is a p-sylow subgroup of H, and that HP/H is a p-sylow subgroup of G/H. (HINT: Compare orders.)

5.16. Prove that a 2-sylow subgroup of A_5 has exactly 5 conjugates.

***5.17.** Let $GL(3, p)$ denote the multiplicative group of all 3×3 nonsingular matrices with entries in Z_p. Show that the set of all matrices of the form

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix},$$

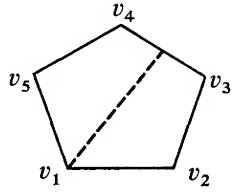
where $a, b, c \in Z_p$, is a p-sylow subgroup of $GL(3, p)$. (HINT: Count.)

SOME APPLICATIONS OF THE SYLOW THEOREMS

We shall now illustrate the power of the Sylow theorems by classifying the groups of small order.

Definition The **dihedral group** D_n, $n \geq 2$, is a group of order $2n$ generated by two elements s and t, which satisfy the relations

$$s^n = 1, \qquad t^2 = 1, \qquad \text{and} \qquad tst = s^{-1}.$$



EXERCISES

5.18. Let A be a regular polygon with vertices $v_1, v_2, \cdots, v_n$, and let G be the set of all rigid motions of A that send vertices into vertices. In particular, let S be a clockwise rotation sending each vertex into the adjacent one, and let T be a reflection of A about the line joining v_1 with the center of A. As usual, multiply in G by first performing one motion and then the other. Prove that $G \cong D_n$.

5.19. Prove that a 2-sylow subgroup of S_5 is isomorphic to D_4.

***5.20.** Show that a 2-sylow subgroup of S_6 is isomorphic to $D_4 \times \sigma(2)$.

5.21. Let $\omega = e^{2\pi i/n}$ be a primitive nth root of unity. Prove that the matrices

$$A = \begin{bmatrix} \omega & 0 \\ 0 & \omega^{-1} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

generate a subgroup of $GL(2, \mathbf{C})$ isomorphic to D_n.

5.22. Show that the set of all matrices

$$\begin{bmatrix} \pm 1 & a \\ 0 & 1 \end{bmatrix}, \qquad a \in Z_n,$$

is a multiplicative group isomorphic to D_n.

5.23. What is the center of D_n? (HINT: First show that every element in D_n has a factorization $s^i t^j$.)

5.24. Let G be a finite group and let $a, b \in G$ be distinct elements of order 2. Prove that $\langle a, b \rangle \cong D_n$ for some n.

5.25. D_4 is not generated by its elements of largest order (compare Exercise 4.16).

5.26. Let G be a finite group with normal subgroups H and K. If $G/H \cong G/K$, is $H \cong K$?

5.27. How many bracelets are there having n beads each of which may be painted any one of q colors? (HINT: See Exercise 3.55; the group D_n is acting.)

Theorem 5.12 *Let p be an odd prime. Any group G of order $2p$ is either cyclic or dihedral.*

Proof By Cauchy's theorem, G contains an element s of order p and an element t of order 2. If $H = \langle s \rangle$, then $H \lhd G$, since it has index 2. Therefore, $tst = s^i$ for some i. Now $s = t^2 s t^2 = t(tst)t = ts^i t =$

s^{i^2}. Hence, $i^2 \equiv 1 \pmod{p}$, so that, since p is prime, $i \equiv \pm 1 \pmod{p}$. Thus, $tst = s$ or $tst = s^{-1}$. In the first case, s and t commute, G is abelian, and so $G \cong \sigma(p) \times \sigma(2) \cong \sigma(2p)$. In the second case, we have $G \cong D_p$. ∎

Theorem 5.13 *Let $|G| = pq$, where $p > q$ are primes. Then either G is cyclic or G is generated by two elements a and b satisfying the following relations:*

$$b^p = 1; \qquad a^q = 1; \qquad a^{-1}ba = b^r,$$

where $r \not\equiv 1 \pmod{p}$ but $r^q \equiv 1 \pmod{p}$.
The second possibility can occur only if q divides $p - 1$.

Proof G contains an element b of order p; let $S = \langle b \rangle$. Since S is a p-sylow subgroup of G, the number of its conjugates is $1 + up$ for some $u \geq 0$. But $1 + up = [G: N(S)]$ which divides $|G| = pq$. Since $(1 + up, p) = 1$, $1 + up$ divides q. Since $q < p$, $u = 0$ and $S \lhd G$.

Now G contains an element a of order q; let $T = \langle a \rangle$. T is a q-sylow subgroup of G, so that $[G: N(T)] = 1 + kq$ for some $k \geq 0$. As above, $1 + kq$ divides p, so that either $k = 0$ or q divides $p - 1$. If $k = 0$, $T \lhd G$ so that $G \cong S \times T$, by Exercise 5.14. Hence, $G \cong \sigma(p) \times \sigma(q) \cong \sigma(pq)$.

We now assume that T is not normal (and so q divides $p - 1$). Since S is normal, $a^{-1}ba = b^r$; furthermore, we may assume $r \not\equiv 1 \pmod{p}$ lest we return to the abelian case. By induction on j, the reader may prove that $a^{-j}ba^j = b^{r^j}$. In particular, if $j = q$, we have $b = b^{r^q}$ so that $r^q \equiv 1 \pmod{p}$. ∎

Corollary 5.14 *Let $|G| = pq$, where $p > q$ are primes. If q does not divide $p - 1$, then G is cyclic.*

Definition[3] The **quaternions**, denoted Q, is a group of order 8 having two generators a and b that satisfy the relations

$$a^4 = 1; \qquad b^2 = a^2; \qquad b^{-1}ab = a^{-1}.$$

[3] In this definition, as in the definition of D_n, we describe a group by generators and relations. At this point, it is fair to ask whether there *is* any group fitting this description (this question will be answered when we deal with extensions and with free groups). Therefore, it is necessary here to exhibit particular groups (motions, matrices, etc.) that do satisfy the conditions.

EXERCISES

5.28. Prove that A_5 contains no subgroup of order 15.

5.29. Let G be the multiplicative group of all 2×2 nonsingular complex matrices, and let H be the subgroup of G generated by

$$A = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Prove that $H \cong Q$.

5.30. Consider the set $\{\pm 1, \pm i, \pm j, \pm k\}$ with the multiplication rules: $i^2 = j^2 = k^2 = -1; ij = k; jk = i; ki = j; ji = -k; kj = -i; ik = -j$; and the usual rules for multiplying by ± 1. Prove that we have described a group isomorphic to Q.

5.31. What is the center of Q?

5.32. Q contains exactly one subgroup of order 2.

5.33. $Q/Z(Q)$ is abelian; moreover, Q contains no subgroup isomorphic to $Q/Z(Q)$ (compare Exercise 4.29).

5.34. Prove that Q is not isomorphic to D_4.

***5.35.** Every subgroup of Q is normal.

5.36. Let $G = Q \times A \times B$, where Q is the quaternions, A is a (necessarily abelian) group of exponent 2, and B is an abelian group in which every element has odd order. Prove that every subgroup of G is normal. (Nonabelian groups in which every subgroup is normal are called **hamiltonian**, after W. R. Hamilton, who discovered the quaternions; all hamiltonian groups are of the form described above.)

Theorem 5.15 *Q and D_4 are the only nonabelian groups of order* 8.

Proof If G is a nonabelian group of order 8, then G has no element of order 8 and not every element of G has order 2 (Exercise 1.29). Thus, G contains an element a of order 4. Suppose $b \in G$ and $b \notin \langle a \rangle$. Since $\langle a \rangle \lhd G$ (it has index 2) and $G/\langle a \rangle \cong \sigma(2)$, we must have $b^2 \in \langle a \rangle$. If $b^2 = a$ or $b^2 = a^3$, then b has order 8, a contradiction. Hence,

$$b^2 = a^2 \quad \text{or} \quad b^2 = 1.$$

Furthermore, since $\langle a \rangle$ is normal, $b^{-1}ab \in \langle a \rangle$. Thus,

$$b^{-1}ab = a \quad \text{or} \quad b^{-1}ab = a^3$$

(because a^2 has order 2). The first case is ruled out, for then a and b commute and G is abelian. The following possibilities remain:

(i) $a^4 = 1$, $b^2 = a^2$, and $b^{-1}ab = a^3$.
(ii) $a^4 = 1$, $b^2 = 1$, and $b^{-1}ab = a^3$.

Since $a^3 = a^{-1}$, (i) describes the quaternions Q and (ii) describes the dihedral group D_4. ∎

Theorem 5.16 *Every group G of order 12 that is not isomorphic to A_4 contains an element of order 6.*

Proof If B is a 3-sylow subgroup of G, then $B = \langle b \rangle$, where b has order 3. Since B has index 4, Theorem 3.17 gives a homomorphism $\pi: G \to S_4$ whose kernel K is a subgroup of B. Now B has no proper subgroups, so that $K = \{1\}$ or $K = B$. If $K = \{1\}$, then π is one-to-one and G is isomorphic to a subgroup of S_4 of order 12, which must be A_4, by Exercise 3.25. Since $G \ncong A_4$, $K = B$, and so B is normal in G. It follows from Exercise 5.9 that B is the unique 3-sylow subgroup of G. Therefore, there are only two elements in G of order 3, namely, b and b^2.

Let $C(b)$ be the centralizer of b in G. Now $[G: C(b)] = 1$ or 2, for a conjugate of b must have order 3. Therefore $|C(b)| = 12$ or 6; in either case, $C(b)$ contains an element a of order 2. Since a commutes with b, the element ab has order 6. ∎

Corollary 5.17 *If G has order 12 and $G \ncong A_4$, then G has a normal 3-sylow subgroup.*

EXERCISES

5.37. Exhibit all the proper subgroups of S_4; there are allegedly 28 of them. How many isomorphism classes of subgroups are there?

5.38. Prove that $D_3 \cong S_3$.

5.39. Prove that $D_6 \cong S_3 \times \sigma(2)$.

***5.40.** Prove that there are only two nonabelian groups of order 12 that contain a subgroup of order 6. One of these is D_6; the other shall be denoted T, and it has generators a and b satisfying the relations $a^6 = 1$ and $b^2 = a^3 = (ab)^2$. (We shall examine the group T more carefully in Chapter 7.)

***5.41.** If G is a nonabelian group of order p^3, where p is a prime, then $Z(G) = G'$, the commutator subgroup of G.

*5.42. Let p be an odd prime. Prove that there are at most two nonabelian groups of order p^3: One has generators a and b satisfying the relations $a^{p^2} = 1$, $b^p = 1$, and $b^{-1}ab = a^{1+p}$; the other has generators a, b, and c satisfying the relations $c = aba^{-1}b^{-1}$, $ca = ac$, $cb = bc$, and $a^p = b^p = c^p = 1$.

*5.43. Let G be a p-sylow subgroup of $GL(3, p)$ (see Exercise 5.17). If p is an odd prime, then G has exponent p. (HINT: If $A \in G$ and E is the identity matrix, expand $(A - E)^p$ by the binomial theorem.)

*5.44. Let G be a finite group and let P be a p-sylow subgroup of G. If H is a normal subgroup of G containing P, then $P \lhd H$ implies $P \lhd G$.

*5.45. If G is a nonabelian group of square-free order (more generally, if every sylow subgroup of G is cyclic), it is known that G is not simple (we know this in the special case when $|G|$ is even, by Exercise 3.41). Use this fact to prove that if $|G| = p_1p_2 \cdots p_t$, where $p_1 < p_2 < \cdots < p_t$ are primes, then G contains a normal p_t-sylow subgroup.

5.46. Prove that there are no nonabelian simple groups of order less than 60. (Do not use the unproven assertion in Exercise 5.45.)

*5.47. Prove that any simple group G of order 60 is isomorphic to A_5. (HINT: Show that a 2-sylow subgroup of G has exactly 5 conjugates so that G has a subgroup of index 5.) In Chapter 8, we shall exhibit two nonisomorphic finite simple groups having the same order.

Definition A **generalized quaternion group** (or *dicyclic group*) Q_n, $n \geq 3$, is a group of order 2^n having generators a and b and relations:

$$a^{2^{n-2}} = b^2 = (ab)^2.$$

EXERCISES

5.48. Let G be a group of order 2^n, which is generated by elements a and b. If

$$a^{2^{n-1}} = 1, \qquad bab^{-1} = a^{-1}, \qquad \text{and} \qquad b^2 = a^{2^{n-2}},$$

then $G \cong Q_n$.

*5.49. Prove that Q_n has a unique subgroup of order 2, which is $Z(Q_n)$.

5.50. Prove that $Q_n/Z(Q_n) \cong D_{2^{n-1}}$.

TABLE OF GROUPS OF SMALL ORDER

Order	Number of Distinct Groups	Groups
4	2	cyclic and 4-group
6	2	$\sigma(6) = \sigma(2) \times \sigma(3)$, S_3
8	5	3 are abelian, D_4, Q
9	2	abelian
10	2	cyclic and dihedral
12	5	2 are abelian, $\sigma(2) \times S_3$, A_4, T
14	2	cyclic and dihedral
15	1	cyclic (Corollary 5.14)

There are 14 nonisomorphic groups of order 16, so we end our list here. No one knows a formula giving, for each n, the number of nonisomorphic groups of order n.

CHAPTER

6

Normal Series

SOME GALOIS THEORY

We begin this chapter with a brief history of the study of polynomials. It was well known to mathematicians of the Middle Ages, even to the Babylonians, that the roots of a quadratic polynomial

$$x^2 + bx + c$$

are given by the formula

$$x = \tfrac{1}{2}[-b \pm (b^2 - 4c)^{1/2}].$$

Now every cubic polynomial,

$$x^3 + ax^2 + bx + c,$$

can be transformed [by sending x into $x - (a/3)$] into a cubic of the form

$$x^3 + qx + r.$$

In 1515, Scipio del Ferro and Tartaglia independently obtained the formula for these roots:

$$x = \left\{-\frac{r}{2} + \left(\frac{r^2}{4} + \frac{q^3}{27}\right)^{1/2}\right\}^{1/3} + \left\{-\frac{r}{2} - \left(\frac{r^2}{4} + \frac{q^3}{27}\right)^{1/2}\right\}^{1/3},$$

where the cube roots are chosen so that their product is $-(q/3)$. In 1545, L. Ferrari found a similar formula for the roots of the general quartic

polynomial. Abel's proof (1824) of the nonexistence of a similar formula for the roots of the general quintic polynomial thus ended 280 years of searching for a generalization of the work of Scipio, Tartaglia, and Ferrari.

Let

$$f(x) = a_0 + a_1x + \cdots + a_nx^n$$

be a polynomial over the complex numbers $\mathbf{C}$; $f(x) \in \mathbf{C}[x]$. We say that there is a **formula** for a root r of $f(x)$ in case r can be obtained from $a_0, a_1, \cdots, a_n$ by a finite number of applications of the operations of addition, subtraction, multiplication, division, multiplication by rationals, and extraction of roots. We shall now discuss fields in order to express this idea in a more convenient way.

Definition Let F be a subfield of $\mathbf{C}$ and let $\alpha \in \mathbf{C}$. $F(\alpha)$, the field obtained by **adjoining** α to F, is the set of all quotients $f(\alpha)/g(\alpha)$, where $f(\alpha)$ and $g(\alpha)$ are polynomials in α with coefficients in F, and $g(\alpha) \neq 0$.

It is easy to check that $F(\alpha)$ is a subfield of $\mathbf{C}$ containing F and α; indeed, it is the smallest such subfield. Note that $F(\alpha) = F$ if and only if $\alpha \in F$.

The process of adjoining a number to F may be iterated: If $\alpha_1, \alpha_2, \cdots, \alpha_n$ are in $\mathbf{C}$, define $F(\alpha_1, \alpha_2, \cdots, \alpha_n) = K(\alpha_n)$, where $K = F(\alpha_1, \alpha_2, \cdots, \alpha_{n-1})$.

EXERCISES

6.1. Every subfield of $\mathbf{C}$ contains the rational numbers, $\mathbf{Q}$.

6.2. Let F be a subfield of $\mathbf{C}$ and let $\{\alpha_1, \alpha_2, \cdots, \alpha_n\} \subset \mathbf{C}$. Prove that $F(\alpha_1, \cdots, \alpha_n) = F(\alpha_{\pi(1)}, \cdots, \alpha_{\pi(n)})$, where π is a permutation of $\{1, 2, \cdots, n\}$. (HINT: Prove that $F(\alpha_1, \alpha_2, \cdots, \alpha_n)$ is the intersection of all the subfields of $\mathbf{C}$ containing F and $\{\alpha_1, \cdots, \alpha_n\}$.)

If F is any field, a polynomial $f(x) \in F[x]$ of degree n has at most n roots in F. If $F = \mathbf{C}$, the fundamental theorem of algebra says that there are complex numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ (not necessarily distinct) such that

$$f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

For the remainder of this discussion, all fields are assumed to be subfields of the complex numbers $\mathbf{C}$.

Definition Let F be a subfield of $\mathbf{C}$ and let $f(x) \in F[x]$. The **root field** of $f(x)$ over F, $R(f, F)$, is the field $F(\alpha_1, \alpha_2, \cdots, \alpha_n)$, where $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is the set of all roots of $f(x)$.

Note that a polynomial has many root fields: If $f(x)$ has roots $\alpha_1, \cdots, \alpha_m$ and if the coefficients of $f(x)$ lie in fields F and K, then $R(f, F) = F(\alpha_1, \cdots, \alpha_m)$ and $R(f, K) = K(\alpha_1, \cdots, \alpha_m)$.

Definition Let $f(x) \in F[x]$ and $R = R(f, F)$; $f(x)$ is **solvable by radicals** in case there are fields

$$F = K_0 \subset K_1 \subset \cdots \subset K_t$$

with

$$R \subset K_t,$$

where each K_{i+1} is obtained from K_i by adjoining a root of an element in K_i.

When we say that there is a formula for the roots of $f(x)$, we really mean that $f(x)$ is solvable by radicals. Let us illustrate this by considering the quadratic and cubic formulas.

If $f(x) = x^2 + bx + c$, set $F = \mathbf{Q}(b, c)$, where $\mathbf{Q}$ is the rationals; observe that $b^2 - 4c \in F$. Let $\alpha = (b^2 - 4c)^{1/2}$, and set $K_1 = F(\alpha)$. The quadratic formula says that K_1 is the root field of $f(x)$ over F.

If $f(x) = x^3 + qx + r$, set $F = \mathbf{Q}(q, r)$; observe that $r^2/4 + q^3/27 \in F$. Let

$$\alpha = \left(\frac{r^2}{4} + \frac{q^3}{27}\right)^{1/2},$$

and set $K_1 = F(\alpha)$. Observe further that $-r/2 + \alpha \in K_1$. Let

$$\beta = \left(-\frac{r}{2} + \alpha\right)^{1/3},$$

and set $K_2 = K_1(\beta)$. Note that $\gamma = -q/3\beta$ is a cube root of $-(r/2) - \alpha$, so that the cubic formula shows that K_2 is the root field of $f(x)$ over F.

Definition If K is a field, an **automorphism** of K is a one-to-one correspondence $\lambda: K \to K$ such that

$$\lambda(x + y) = \lambda(x) + \lambda(y)$$

and

$$\lambda(xy) = \lambda(x)\lambda(y)$$

for every x and y in K. If F is a subfield of K, we say that λ **fixes** F in case $\lambda(a) = a$ for every $a \in F$.

Lemma 6.1 *Let F be a subfield of K. The set of all automorphisms of K that fix F forms a group (in which the binary operation is composition of functions).*

Proof Trivial. ∎

Definition Let $f(x) \in F[x]$ and let $R = R(f, F)$. The **Galois group** of $f(x)$ over F, denoted $G(R/F)$, is the group of all automorphisms of R fixing F.

Theorem 6.2 *Let $X = \{\alpha_1, \cdots, \alpha_k\}$ be all the roots of a polynomial $f(x) \in F[x]$ and let $R = F(\alpha_1, \cdots, \alpha_k)$ be the root field of $f(x)$ over F. Then the map $\lambda \to \lambda|X$ is an imbedding $G(R/F) \to S_X \cong S_k$.*

Proof Suppose $f(x) = a_0 + a_1x + \cdots + a_nx^n$, where each $a_i \in F$. If $\beta \in R$, then

$$f(\beta) = a_0 + a_1\beta + \cdots + a_n\beta^n \in R$$

and

$$\lambda(f(\beta)) = a_0 + a_1\lambda(\beta) + \cdots + a_n\lambda(\beta)^n = f(\lambda(\beta)),$$

since λ fixes F. In particular, if β is a root of $f(x)$, so is $\lambda(\beta)$. Therefore, $\lambda|X: X \to X$; indeed, $\lambda|X \in S_X$ because λ is one-one and X is finite.

It is easy to show this function $G(R/F) \to S_X$ is a homomorphism. To see it is one-one, observe that if λ fixes F and $\alpha_1, \cdots, \alpha_k$, then λ fixes $R = F(\alpha_1, \cdots, \alpha_k)$. ∎

We now present some field theory to establish a connection between the solvability of a polynomial by radicals and its Galois group.

Let K be a field with subfield F. Clearly K satisfies the axioms for a vector space over F if we take as scalar multiplication the given multiplication between elements of F and K. In particular, one may ask about the dimension of K over F, denoted $[K: F]$.

EXERCISES

****6.3.** Let $p(x) \in F[x]$ be an irreducible polynomial of degree n. If $\alpha \in \mathbf{C}$ is a root of $p(x)$, prove that $\{1, \alpha, \alpha^2, \cdots, \alpha^{n-1}\}$ is a basis of $F(\alpha)$ viewed as a vector space over F. Conclude that $[F(\alpha): F] = n$.

**6.4. Let $F \subset K \subset R$ be fields, where $[K: F]$ and $[R: K]$ are finite. Then

$$[R: F] = [R: K][K: F].$$

(HINT: If $\{\alpha_1, \cdots, \alpha_r\}$ is a basis of K over F and $\{\beta_1, \cdots, \beta_s\}$ is a basis of R over K, then the set of rs elements of the form $\alpha_i\beta_j$ is a basis of R over F.)

**6.5. Let $F \subset K \subset R$, where R and K are each root fields over F. Let λ be an automorphism of R fixing F; prove that $\lambda|K \in G(K/F)$.

Lemma 6.3 *Let $\lambda: F \to F'$ be an isomorphism of fields, and let $p(x) \in F[x]$ be irreducible with root $\alpha \in \mathbf{C}$. Then there is an irreducible polynomial $q(x) \in F'[x]$ with root $\beta \in \mathbf{C}$ and an isomorphism*

$$\lambda^*: F(\alpha) \to F'(\beta)$$

with $\lambda^*|F = \lambda$ and $\lambda^*(\alpha) = \beta$.

Proof The map $\varphi: F[x] \to F'[x]$ defined by

$$b_0 + b_1x + \cdots + b_mx^m \to \lambda(b_0) + \lambda(b_1)x + \cdots + \lambda(b_m)x^m$$

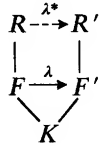
is an isomorphism of rings. Define $q(x) = \varphi(p(x))$, which is irreducible since $p(x)$ is; let $\beta \in \mathbf{C}$ be any root of $q(x)$.

Since $p(x)$ is irreducible, Exercise 6.3 says $\{1, \alpha, \cdots, \alpha^{n-1}\}$ is a basis of $F(\alpha)$ over F, where $n =$ degree $p(x)$. Define $\lambda^*: F(\alpha) \to F'(\beta)$ by

$$\lambda^*(b_0 + d_1\alpha + \cdots + b_{n-1}\alpha^{n-1}) = \lambda(b_0) + \lambda(b_1)\beta + \cdots + \lambda(b_{n-1})\beta^{n-1}.$$

It is easy to check that λ^* is a homomorphism with $\lambda^*|F = \lambda$ and $\lambda^*(\alpha) = \beta$. To see that λ^* is an isomorphism, construct its inverse in the same manner, beginning with $\lambda^{-1}: F' \to F$. ∎

Lemma 6.4 *Let $f(x) \in K[x]$, where K is a subfield of F and of F'; let $R = R(f, F)$ and $R' = R(f, F')$. If $\lambda: F \to F'$ is an isomorphism fixing K, then there is an isomorphism $\lambda^*: R \to R'$ with $\lambda^*|F = \lambda$.*



Proof The proof is by induction on $d = [R: F]$. If $d = 1$, then $R = F$ and $R' = F'$, and there is nothing to prove. If $d > 1$, then $R \neq F$, so there is some root α of $f(x)$ not lying in F (of course $\alpha \in R$, by definition of root field). Now α is a root of some irreducible factor $p(x)$ of $f(x)$; since $\alpha \notin F$, if $k =$ degree $p(x)$, then $k > 1$. By Lemma 6.3, there is $\beta \in \mathbf{C}$ and an isomorphism $\lambda_1: F(\alpha) \to F'(\beta)$ that extends λ. In particular, λ_1 fixes K, so that β is a root of $f(x)$ and $F'(\beta) \subset R'$. By Exercise 6.4, $[R: F(\alpha)] = d/k < d$. Furthermore, R arises from $F(\alpha)$ by adjoining all the roots of $f(x)$, so that R is the root field of $f(x)$ over $F(\alpha)$; similarly, R' is the root field of $f(x)$ over $F'(\beta)$. With all the inductive hypotheses being verified, λ_1, a fortiori λ, can be extended to an isomorphism $\lambda^*: R \to R'$. ∎

Lemma 6.5 *Let $F \subset K \subset R$, where R and K are each root fields over F. Then $G(R/K) \lhd G(R/F)$ and*

$$G(R/F)/G(R/K) \cong G(K/F).$$

Proof Define $\Phi: G(R/F) \to G(K/F)$ by

$$\Phi(\lambda) = \lambda|K.$$

By Exercise 6.5, Φ is a well-defined function that is easily seen to be a homomorphism. The kernel of Φ is the set of all automorphisms of R that fix K, i.e., $G(R/K)$. We claim that Φ is onto. Suppose $\lambda \in G(K/F)$, i.e., λ is an automorphism of K that fixes F. Since R is a root field of F, Lemma 6.4 says λ can be extended to an automorphism λ^* of R. Therefore, $\lambda^* \in G(R/F)$ and

$$\Phi(\lambda^*) = \lambda^*|K = \lambda.$$

The first isomorphism theorem completes the proof. ∎

From this point, F is not only assumed to be a subfield of $\mathbf{C}$, *it is also assumed to contain all roots of unity.*

EXERCISES

****6.6.** Let $R = F(\alpha)$, where α is a root of $x^k - a$, $a \in F$. Prove that there exist intermediate fields

$$F = K_0 \subset K_1 \subset \cdots \subset K_t = R,$$

where $K_{i+1} = K_i(\alpha_i)$, α_i is a root of $x^{n(i)} - a_i$, $a_i \in K_i$, and $n(i)$ is prime.

6.7. If p is a prime and $a \in F$, then $x^p - a$ either has a root in F or is irreducible over F.

[NOTE: If $f(x)$ has a root $a \in F$, then $f(x)$ is not irreducible, for it is divisible by $x - a$. The converse is false: If $f(x)$ factors, it need not have a root in F. For example, if F is the reals, then $f(x) = (x^2 + 1)(x^2 + 1)$ factors, but it has no real roots.]

****6.8.** Let α be a root of $x^p - a$, $a \in F$, and p be a prime. Prove that $G(F(\alpha)/F)$ is a cyclic group of order p or 1.

We summarize this investigation in the next theorem.

Theorem 6.6 *Assume F is a subfield of $\mathbf{C}$ that contains all roots of unity and R is the root field of $f(x) \in F[x]$. If $f(x)$ is solvable by radicals, then there exists subgroups G_i of $G = G(R/F)$ such that:*

(i) $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$.
(ii) $G_{i+1} \lhd G_i$, *for all* i.
(iii) G_i/G_{i+1} *is cyclic of prime order.*

Proof Define $G_i = G(R/K_i)$, where the K_i are as in Exercise 6.6. We only remind the reader of previous work: (i) follows immediately from the definition of Galois group; (ii) follows from Lemma 6.5; (iii) follows from Exercises 6.6 and 6.8 and Lemma 6.5. ∎

The converse of this theorem is also true, but we shall not prove it here. We have been led to the following concept.

Definition A **normal series** of G is a chain of subgroups

$$(*) \qquad G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$$

in which $G_{i+1} \lhd G_i$ for all i. The **factor groups** of (*) are the groups G_i/G_{i+1} for $i = 0, 1, \cdots, n - 1$; the **length** of the series (*) is the number of strict inclusions.

Alternatively, the length of the series is the number of factor groups with more than one element. Note that the factor groups are the only quotient groups we can always form, for we have seen (Exercise 3.33) that normality need not be transitive.

Definition A finite group G is **solvable** in case it has a normal series whose factor groups are cyclic of prime order.

In this terminology, Theorem 6.6 and its converse say that a polynomial $f(x)$ is solvable by radicals if and only if its Galois group $G(R/F)$ is a solvable group. Now Abel proved that there exists a polynomial $f(x)$ of degree 5 whose Galois group[1] is S_5. Since S_5 is not a solvable group (as we shall soon see), this polynomial is not solvable by radicals. Thus, the classical problem of the determination of the roots of a polynomial led inevitably to groups and to normal series, and there its solution (or nonsolution) lies. (We refer the interested reader to E. Artin's monograph, *Galois Theory*, and B. L. van der Waerden's *Modern Algebra*, Vol. I (see Bibliography), for a more thorough discussion.)

The work of Abel and Galois has not only enriched the study of polynomials and fields, it has also contributed a new point of view to the study of groups. Let us give a brief review of what we have learned so far. Our first results arose from examining properties of a single subgroup via Lagrange's theorem. The second, deeper set of results arose from examining properties of a family of subgroups via the Sylow theorems (Theorems 5.9 and 5.10); this family of subgroups consists of the conjugates of a single subgroup, and so each member has the same order. Normal series will give results by allowing us to examine a family of subgroups of distinct orders, thus providing an opening wedge for an inductive proof.

THE JORDAN-HÖLDER THEOREM

Definition Let

$$(*) \qquad G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$$

and

$$(**) \qquad G = H_0 \supset H_1 \supset \cdots \supset H_m = \{1\}$$

be normal series of G; (**) is a **refinement** of (*) in case (**) = (*) or (**) is obtained from (*) by the insertion of subgroups.

The length of a refinement is thus at least as great as the length of the original series.

Definition A normal series

$$G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$$

[1] Galois outlined Galois theory in 1830 (six years after Abel proved the unsolvability of the quintic), when he was only 19; he was killed in a duel two years later.

is a **composition series** in case each G_{i+1} is a maximal normal subgroup of G_i.

EXERCISES

6.9. A normal series is a composition series if and only if each of its factor groups is simple or $\{1\}$.

6.10. A composition series is a normal series of maximal length.

6.11. Every finite group has a composition series.

6.12. An abelian group has a composition series if and only if it is finite.

6.13. If G is a finite group having a normal series with factor groups $H_1, \cdots, H_m$, then $|G| = \prod_i |H_i|$.

6.14. Give an example of an infinite group with a composition series.

6.15. A subgroup H is **subnormal** in G in case there is a normal series from G to H. Prove that a group G has a composition series if and only if G has both chain conditions on subnormal subgroups.

Consider the group $G = \sigma(30)$ with generator x; we write two composition series for G (normality is automatic, since G is abelian):

$$(*) \qquad G \supset \langle x^5 \rangle \supset \langle x^{10} \rangle \supset \{1\};$$

$$(**) \qquad G \supset \langle x^2 \rangle \supset \langle x^6 \rangle \supset \{1\}.$$

The factor groups of (*) are $G/\langle x^5 \rangle$, $\langle x^5 \rangle/\langle x^{10} \rangle$, and $\langle x^{10} \rangle$, i.e., $\sigma(5)$, $\sigma(2)$, and $\sigma(3)$. The factor groups of (**) are $G/\langle x^2 \rangle$, $\langle x^2 \rangle/\langle x^6 \rangle$, and $\langle x^6 \rangle$, i.e., $\sigma(2)$, $\sigma(3)$, $\sigma(5)$. In this case, these two composition series of G have the same length, and the factor groups can be "paired isomorphically" after rearranging them. We give a name to this phenomenon.

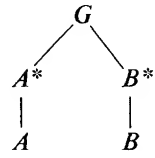
Definition Two normal series (*) and (**) of G are **equivalent** if there is a one-to-one correspondence between the factor groups of (*) and (**) such that corresponding factor groups are isomorphic.

Observe that equivalent normal series have the same length.

The two composition series for $\sigma(30)$ exhibited above are equivalent; the amazing fact is that this is true for every (possibly infinite) group that has a composition series!

The next result, a generalization of the second isomorphism theorem, will be used in the proof of equivalence of composition series.

Theorem 6.7 (Zassenhaus Lemma) *Let A, A^*, B and B^* be four subgroups of a group G with $A \lhd A^*$ and $B \lhd B^*$.*



Then

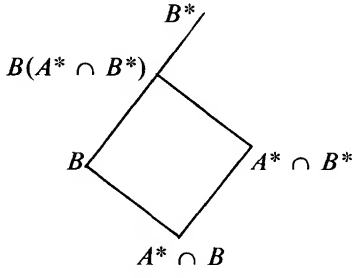
$$A(A^* \cap B) \lhd A(A^* \cap B^*),$$

$$B(A \cap B^*) \lhd B(A^* \cap B^*),$$

and there is an isomorphism

$$\frac{A(A^* \cap B^*)}{A(A^* \cap B)} \cong \frac{B(A^* \cap B^*)}{B(A \cap B^*)}.$$

Proof Let D be the subset $(A^* \cap B)(A \cap B^*)$. Now $B \lhd B^*$ implies $BS = SB$ for any nonempty subset S of B^*. It follows from Exercise 2.53 that $B(A^* \cap B^*)$ is a subgroup of B^* and that



$B \lhd B(A^* \cap B^*)$ (for B is normal in the larger group B^*). By the second isomorphism theorem, $A^* \cap B \lhd A^* \cap B^*$. In a similar manner, one sees that $A \cap B^* \lhd A^* \cap B^*$. Therefore Theorem 2.13 and Exercise 2.42 show that the subset $D = (A^* \cap B)(A \cap B^*)$ is a normal subgroup of $A^* \cap B^*$.

If $x \in B(A^* \cap B^*)$, then $x = bc$ for $b \in B$ and $c \in A^* \cap B^*$; define f: $B(A^* \cap B^*) \to (A^* \cap B^*)/D$ by $f(x) = Dc$. It is straightforward to check that f is a well-defined homomorphism of

$B(A^* \cap B^*)$ onto $(A^* \cap B^*)/D$ with kernel $B(A \cap B^*)$. By the first isomorphism theorem, $B(A \cap B^*) \lhd B(A^* \cap B^*)$ and

$$\frac{B(A^* \cap B^*)}{B(A \cap B^*)} \cong \frac{A^* \cap B^*}{D}.$$

If we interchange the letters A and B, then the symmetry of the hypothesis gives a homomorphism $A(A^* \cap B^*) \to (A^* \cap B^*)/D$; hence $A(A^* \cap B) \lhd A(A^* \cap B^*)$ and the quotient is also isomorphic to $(A^* \cap B^*)/D$. Therefore the two quotient groups of the theorem are isomorphic. ∎

We remark that the Zassenhaus lemma generalizes the second isomorphism theorem: set $A = T$, $A^* = G$, $B = S \cap T$, and $B^* = S$, where S and T are subgroups of G with $T \lhd G$.

Theorem 6.8 (Schreier) *Any two normal series of an arbitrary group G have refinements that are equivalent.*

Proof Let

$$(*) \qquad G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$$

and

$$(**) \qquad G = H_0 \supset H_1 \supset \cdots \supset H_m = \{1\}$$

be normal series. In the first series, between each G_i and G_{i+1} insert the groups $G_{i+1}(G_i \cap H_j)$, where $0 \leq j \leq m$. These are the "numerators" of the groups in the Zassenhaus lemma. The normality assertions of that lemma say that this is a normal series with mn (not necessarily strict) inclusions. In the second series, between each H_j and H_{j+1} insert the groups $H_{j+1}(H_j \cap G_i)$, where $0 \leq i \leq n$; this is also a normal series with mn inclusions. We now have normal series

$$\text{(i)} \qquad \cdots \supset G_{i+1}(G_i \cap H_j) \supset G_{i+1}(G_i \cap H_{j+1}) \supset \cdots$$

and

$$\text{(ii)} \qquad \cdots \supset H_{j+1}(H_j \cap G_i) \supset H_{j+1}(H_j \cap G_{i+1}) \supset \cdots$$

These are common refinements: Series (i) visibly refines series (*) (take $j = m$); series (i) also refines series (**), because, given j, there is an index i with $G_i \supset H_j \supset G_{i+1}$, so that $H_j = G_{i+1}(G_i \cap H_j)$. Similarly, series (ii) refines both (*) and (**). The isomorphism of Theorem 6.7 with the four subgroups G_{i+1}, G_i, H_{j+1}, H_j is precisely what is needed to complete the proof. ∎

Theorem 6.9 (Jordan-Hölder)[2] *Any two composition series of a group G are equivalent.*

Proof Since composition series are normal series, any two composition series of G have equivalent refinements. But a composition series is a normal series of maximal length, so it admits no nontrivial refinement (i.e., a refinement merely repeats some of the terms). Thus, two composition series are already equivalent. ▮

Definition The factor groups of a composition series of G are called the **composition factors** of G.

The Jordan-Hölder theorem may be regarded as a kind of unique factorization theorem. In a sense to be made explicit in the next chapter, a group G is a "product" of its composition factors. Let us support this viewpoint by using the Jordan-Hölder theorem to prove the fundamental theorem of arithmetic (Appendix V): Every integer $n > 1$ has a unique factorization into primes. Let $n = p_1 p_2 \cdots p_m$, where the p_i are (not necessarily distinct) primes. If $G = \sigma(n)$ has generator x, then

$$G = \langle x \rangle \supset \langle x^{n/p_1} \rangle \supset \langle x^{n/p_1 p_2} \rangle \supset \cdots \supset \langle x^{p_m} \rangle \supset \{1\}$$

is a normal series. Since the factor groups $\sigma(p_1), \sigma(p_2), \cdots, \sigma(p_m)$ have prime order, this is a composition series. The Jordan-Hölder theorem says that these factor groups, hence their orders, depend only on $G = \sigma(n)$ and not on the choice of composition series.

The proof given above is valid, *mutatis mutandis*, for R-modules. In particular, if R is a field, the Jordan-Hölder theorem gives another proof that any two bases of a finite dimensional vector space have the same size, i.e., that dimension is well defined, for R itself is the only simple R-module. Generalizations of Jordan-Hölder to more general situations, e.g., operator-groups, lattices, can be found in the book of Zassenhaus.

EXERCISES

****6.16.** Let G and H be groups having composition series. If there are normal series of G and H having the same set of factor groups, then G and H have the same composition factors.

6.17. The normal series of S_4 displayed in Exercise 3.33 is a composition series.

6.18. S_n is solvable for $n \leq 4$.

[2] Jordan proved that the orders of the factor groups of a composition series of G depend only on G; Hölder proved that the factor groups depend only on G.

A finite group G is called solvable if it has a normal series with factor groups of prime order. Clearly such a series is a composition series. To test whether a particular group G is solvable, choose a composition series and check its factor groups. It is conceivable, without Jordan-Hölder, that one might have to look at all composition series of a group G before concluding G is not solvable.

Theorem 6.10 *S_n is not solvable for $n \geq 5$.*

Proof A normal series for S_n is

$$S_n \supset A_n \supset \{1\}.$$

Since the factor groups are $\sigma(2)$ and A_n and since A_n is simple for $n \geq 5$, $S_n \supset A_n \supset \{1\}$ is a composition series. By the Jordan-Hölder theorem, S_n is not solvable. ∎

We confess that we really don't need Jordan-Hölder to prove Theorem 6.10, for it is easy to prove that for $n \geq 5$, S_n has a unique composition series (see Exercise 3.32).

SOLVABLE GROUPS

Solvable groups have been defined in connection with Galois theory. Since these groups form a large class of groups of purely group-theoretical interest as well, we give another definition of solvability which is more convenient to work with and which is equivalent to our earlier definition when G is finite.

Definition A group G is **solvable** if it has a normal series with abelian factor groups. Such a normal series is called a **solvable series** of G.

EXERCISES

****6.19.** Any refinement of a solvable series is a solvable series.

6.20. A finite group is solvable (definition just given) if and only if it has a composition series with cyclic factor groups of prime order (earlier definition).

6.21. A solvable group with a composition series is finite.

6.22. Let $H \lhd G$, where G has a composition series. Prove that G has a composition series one of whose terms is H.

Let us see how to manufacture solvable groups.

Theorem 6.11 *Any subgroup H of a solvable group G is solvable.*

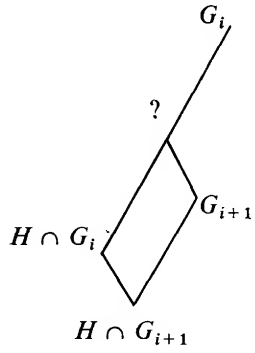
Proof Let $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$ be a solvable series for G; we claim that

$$H = H_0 \supset (H \cap G_1) \supset \cdots \supset (H \cap G_n) = \{1\}$$

is a solvable series for H. By the second isomorphism theorem,

$$H \cap G_{i+1} = (H \cap G_i) \cap G_{i+1} \lhd H \cap G_i,$$

so that we do have a normal series. Moreover, we have the diagram



where $? = G_{i+1}(H \cap G_i)$. Therefore,

$$\frac{H \cap G_i}{H \cap G_{i+1}} \cong \frac{?}{G_{i+1}} \subset \frac{G_i}{G_{i+1}}.$$

Since G_i/G_{i+1} is abelian, so is its subgroup $?/G_{i+1}$, as desired. ∎

Theorem 6.12 *If G is solvable and $H \lhd G$, then G/H is solvable.*

Proof Let $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$ be a solvable series. Now $G \supset H \supset \{1\}$ is a normal series, so that, by Theorem 6.8, these two series have equivalent refinements. Therefore (Exercise 6.19), there is a solvable series

$$G = K_0 \supset K_1 \supset \cdots \supset H \supset \cdots \supset \{1\}.$$

Consider the series

$$G/H \supset K_1/H \supset K_2/H \supset \cdots \supset H/H = \{1\}.$$

By the third isomorphism theorem, this is a normal series and $(K_i/H)/(K_{i+1}/H) \cong K_i/K_{i+1}$, which is abelian. Hence, G/H is solvable. ▮

(Another proof of this theorem, to be described shortly, does not use Theorem 6.8).

Theorem 6.13 *Let $H \lhd G$. If both H and G/H are solvable, then G is solvable.*

Proof Let

$$G/H \supset K_1^* \supset K_2^* \supset \cdots \supset \{1\}$$

be a solvable series. By the correspondence theorem, we can construct the beginning of a solvable series from G to H, i.e., there are subgroups K_i with $K_{i+1} \lhd K_i$, K_i/K_{i+1} abelian, and

$$G = G_0 \supset K_1 \supset K_2 \supset \cdots \supset H.$$

Since H is solvable, it has a solvable series. If we hook these two series together at H, we obtain a solvable series for G. ▮

Corollary 6.14 *If H and K are solvable, then $H \times K$ is solvable.*

Proof If $G = H \times K$, then $H \lhd G$ and $G/H \cong K$, so that G is solvable, by Theorem 6.13. ▮

Corollary 6.15 *Every finite p-group G is solvable.*

Proof We perform an induction on $|G|$. By Theorem 5.4, $|Z(G)| \neq 1$. Therefore, $G/Z(G)$ is a p-group of order $< |G|$, and so is solvable, by induction. Because every abelian group is solvable, $Z(G)$ is solvable. By Theorem 6.13, G is solvable. ▮

EXERCISES

6.23. If S and T are solvable subgroups of G with $S \lhd G$, then ST is a solvable subgroup of G.

6.24. Any group of order p^2q is solvable, where p and q are primes. (Burnside has proved, using representation theory, i.e., the study of homomorphisms of groups into groups of matrices, that any group of order $p^m q^n$ is solvable, where p and q are primes.)

6.25. Any group of square-free order is solvable. (HINT: Use Exercise 5.45.)

6.26. The dihedral groups D_n are solvable.

***6.27.** The following two statements are equivalent: (1) Every group of odd order is solvable; (2) every finite nonabelian simple group has even order. (In a deep paper in 1963, W. Feit and J. Thompson proved the truth of these statements.)

Another approach to solvable groups is with commutator subgroups: That we are dealing with abelian quotient groups suggests this approach at once.

Definition Define the **higher commutator subgroups** of G inductively:

$$G^{(0)} = G; \qquad G^{(i+1)} = G^{(i)\prime},$$

i.e., $G^{(i+1)}$ is the commutator subgroup of $G^{(i)}$. The series of higher commutator subgroups is called the **derived series** of G.

We also want the following terminology.

Definition A subgroup H of G is **characteristic** in G in case $f(H) \subset H$ for every isomorphism $f\colon G \to G$; H is **fully invariant** in G in case $f(H) \subset H$ for every homomorphism $f\colon G \to G$.

Every fully invariant subgroup is characteristic, and every characteristic subgroup is normal.

EXERCISES

6.28. Let $G \neq \{1\}$ be a finite group. If G is solvable, then G contains a normal abelian subgroup $H \neq \{1\}$; if G is not solvable, then G contains a normal subgroup $H \neq \{1\}$ such that $H = H'$.

****6.29.** If $H \subset N \lhd G$ and H is characteristic in N, then $H \lhd G$. (Compare Exercise 5.44.)

6.30. For any group G, $Z(G)$ is characteristic (it need not be a fully invariant subgroup).

****6.31.** If G is an abelian p-group, then $G[p]$, the set of all elements in G of order $\leq p$, is fully invariant in G.

****6.32.** Every normal p-sylow subgroup of a finite group G is fully invariant in G.

6.33. Every $G^{(i)}$ is a fully invariant subgroup of G.

Lemma 6.16 *If* $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$ *is a solvable series, then* $G_i \supset G^{(i)}$ *for all* i.

Proof We prove the lemma by induction on i. If $i = 0$, $G_i = G^{(i)} = G$. Suppose now that $G_i \supset G^{(i)}$; then $G_i' \supset G^{(i)'} = G^{(i+1)}$. Since G_i/G_{i+1} is abelian, $G_{i+1} \supset G_i'$, by Exercise 2.50. Therefore, $G_{i+1} \supset G^{(i+1)}$, as desired. ∎

Theorem 6.17 *A group* G *is solvable if and only if* $G^{(n)} = \{1\}$ *for some integer* n.

Proof Let $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$ be a solvable series. By the lemma, $G_n \supset G^{(n)}$, and so $G^{(n)} = \{1\}$.

If $G^{(n)} = \{1\}$ for some n, then the series

$$G = G^{(0)} \supset G' \supset \cdots \supset G^{(n)} = \{1\}$$

is a solvable series for G. ∎

The reader should supply alternative proofs of Theorems 6.11 and 6.12 using this new characterization of solvability.

A THEOREM OF P. HALL

The main result of this section is a generalization of the Sylow theorems that holds for (and, in fact, characterizes) finite solvable groups.

Definition A **minimal normal subgroup** N of G is a normal subgroup $\neq \{1\}$ that contains no proper subgroup that is normal in G.

Lemma 6.18 *Every minimal normal subgroup* N *of a finite solvable group* G *is an elementary abelian group.*

Proof N' is a fully invariant subgroup of N so that $N' \lhd G$, by Exercise 6.29. Since N is a minimal normal subgroup of G, either $N' = N$ or $N' = \{1\}$. Since N is a subgroup of a solvable group, N is itself solvable. Therefore, $N \neq N'$ and $N' = \{1\}$, i.e., N is abelian.

Because N is finite, it is the direct product of its sylow subgroups. By Exercises 6.32 and 6.29, N is a p-group. Finally, Exercises 6.31 and 6.29 imply that N is elementary. ∎

This lemma is a special case of a more general result: A minimal normal subgroup of a (not necessarily solvable) finite group is a direct product of isomorphic simple groups.

Theorem 6.19 (P. Hall) *Let G be a solvable group of order ab, where $(a, b) = 1$. Then G contains at least one subgroup of order a, and any two such are conjugate.*

Proof The proof proceeds by induction on $|G|$, the theorem holding when $|G| = 1$.

CASE (i) G contains a normal subgroup H of order a_1b_1, where a_1 divides a, b_1 divides b, and $b_1 < b$.

G/H is a solvable group of order $(a/a_1) \cdot (b/b_1)$ and so contains a subgroup $\hat{A}/H$ of order a/a_1. The subgroup $\hat{A}$ of G has order $ab_1 < ab$. Since $\hat{A}$ is also solvable, it contains a subgroup of order a, as desired.

Suppose A and A_1 are subgroups of G of order a. Let us compute $k = |AH|$. By Lagrange's theorem, $k \mid ab = |G|$ while $a_1b_1 \mid k$ and $a|k$. By the second isomorphism theorem (actually, by the product formula), $k \mid aa_1b_1 = |A|\,|H|$. Since $(a, b) = 1$, it follows that $k = |AH| = ab_1$. In a similar manner, $|A_1H| = ab_1$ also. Thus AH/H and A_1H/H are subgroups of G/H of order a/a_1; by induction, these subgroups are conjugate (say, by $\bar{x} \in G/H$). If $x \in G$ goes into $\bar{x}$ under the natural map, it is quickly checked that $xAHx^{-1} = A_1H$. Therefore, xAx^{-1} and A_1 are subgroups of A_1H of order a and so are conjugate, by induction. This completes case (i).

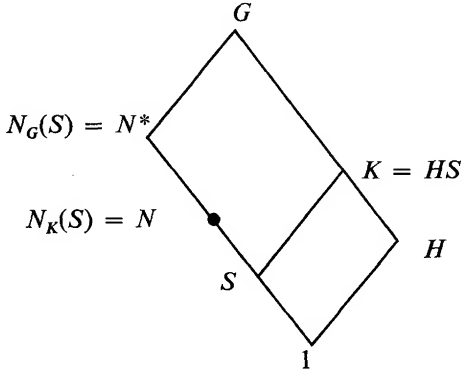
If there is some proper normal subgroup of G whose order is not divisible by b, the theorem is proved. We may therefore assume that b divides $|H|$ for every proper normal subgroup H. If H is a minimal normal subgroup, however, $|H| = p^m$ for some prime p, by Lemma 6.18. We conclude that $b = p^m$. Thus, H is a p-sylow subgroup of G; that H is normal implies that H is the unique such subgroup. The problem has now been reduced to the following case.

CASE (ii) G has a unique minimal normal subgroup H (and H is an abelian group of order p^m and index prime to p).

We finish the proof now, but note that this case follows immediately from the Schur-Zassenhaus lemma, which will be proved in the next chapter.

Observe that since G is finite, every normal subgroup of G contains a minimal normal subgroup of G; in the present case, every proper normal subgroup of G must contain H.

Let K/H be a minimal normal subgroup of G/H. By Lemma 6.18, $|K/H| = q^n$, so that $|K| = p^m q^n$. Let S be a q-sylow subgroup of K and let N^* be the normalizer of S in G. We shall show that $|N^*| = a$.



Observe that $HS \subset K$, and since $H \cap S = \{1\}$, $|HS| = |K|$; therefore, $K = HS$.

Since $K \lhd G$, every conjugate of S in G lies in K, so that all these subgroups are already conjugate in K. Thus, if N is the normalizer of S in K and c is the number of conjugates of S in G, we have

$$[G : N^*] = c = [K : N].$$

Now $K = HN$ (for $S \subset N \subset K$ and $K = HS$), so that

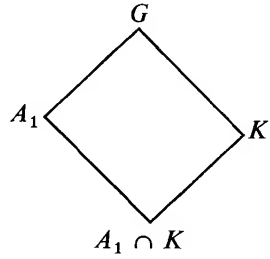
$$c = [G : N^*] = [HN : N] = [H : H \cap N].$$

If we can show that $H \cap N = \{1\}$, then $c = |H| = p^m$, and so $|N^*| = a$. We do this in two stages: (1) $H \cap N \subset Z(K)$; (2) $Z(K) = \{1\}$.

Let $x \in H \cap N$. If $k \in K$, then $k = hs$, where $h \in H$ and $s \in S$. Since $x \in H$, x commutes with h, for H is abelian. Hence, we need show only that x commutes with s. Now $(x^{-1}s^{-1}x)s \in S$, since $x \in N = N_K(S)$; $x^{-1}(s^{-1}xs) \in H$, since H is normal. Therefore, $x^{-1}s^{-1}xs \in S \cap H = \{1\}$, and x commutes with s, as desired.

Finally, $Z(K)$ is a characteristic subgroup of K and $K \lhd G$, so that $Z(K) \lhd G$. If $Z(K) \neq \{1\}$, then $Z(K)$ contains a minimal normal subgroup of G; thus $H \subset Z(K)$. This, together with $K = HS$, tells us that $S \lhd K$. By Exercise 6.38, $S \lhd G$ and so S, too, contains H, a contradiction.

Suppose A_1 is another subgroup of G of order a. Since $|A_1K|$ is divisible by a and by $|K| = p^m q^n$, $|A_1K| = |G| = ab$ and $A_1K = G$.



Therefore,

$$G/K = A_1K/K \cong A_1/A_1 \cap K$$

implies $|A_1 \cap K| = q^n$. By the Sylow theorem, $A_1 \cap K$ is conjugate to S. Since conjugate subgroups have conjugate normalizers, $N^* = N_G(S)$ is conjugate to $N_G(A_1 \cap K)$. Therefore, $a = |N^*| = |N_G(A_1 \cap K)|$. Since $A_1 \cap K \lhd A_1$, we have $A_1 \subset N_G(A_1 \cap K)$. It follows that $A_1 = N_G(A_1 \cap K)$, for both have order a. Hence N^* is conjugate to A_1. ∎

Let us compare Hall's theorem to the Sylow theorems. Suppose $|G| = p^m a$, where $(p^m, a) = 1$. The Sylow theorems say that G has at least one subgroup of order p^m and that any two such are conjugate; Hall's theorem says that if G is solvable, G contains at least one subgroup of order a, and any two such are conjugate. Subgroups whose order and index are relatively prime are thus called **Hall subgroups.**

Definition Let G be a group of order $p^m a$, where $(p^m, a) = 1$. A subgroup of G of order a is a ***p*-complement** of G.

Hall's theorem tells us that every finite solvable group has a p-complement, for every p. The converse of this statement is also true: If G is a finite group containing a p-complement for every p, then G is solvable. Notice that this converse gives an immediate proof of Burnside's theorem: If $|G| = p^m q^n$, then G is solvable. In this case, the p- and q-complements are just q- and p-sylow subgroups. Unfortunately, the proof of the converse (also due to P. Hall) makes use of Burnside's theorem, which we cannot prove here.

CENTRAL SERIES AND NILPOTENT GROUPS

The Sylow theorems show that knowledge of p-groups gives information about arbitrary finite groups. Moreover, p-groups have a rich supply of normal subgroups, and hence admit many homomorphisms; this suggests that normal series might be a powerful tool in their study.

Definition Let H and K be subgroups of G. Then

$$[H, K] = \langle h^{-1}k^{-1}hk : h \in H, k \in K \rangle.$$

EXERCISES

6.34. For every two subgroups H and K of G,

$$[H, K] = [K, H].$$

6.35. $G' = [G, G]$ and, for all i, $G^{(i+1)} = [G^{(i)}, G^{(i)}]$.

****6.36.** A subgroup $K \subset N_G(H)$ if and only if $[H, K] \subset H$. One says that K **normalizes** H in this case.

6.37. $[H, K] = \{1\}$ if and only if each $k \in K$ commutes with every element of H. One says that K **centralizes** H in this case.

Definition If H is a subgroup of G, the **centralizer** of H in G is

$$C_G(H) = \{a \in G : [a, h] = 1 \text{ for all } h \in H\}.$$

EXERCISES

****6.38.** Let $K \lhd G$ and $K \subset H \subset G$. Then $[H, G] \subset K$ if and only if $H/K \subset Z(G/K)$.

****6.39.** Let $f: G \to H$ be a homomorphism onto. If $A \subset Z(G)$, then $f(A) \subset Z(H)$.

Definition Define a chain of subgroups $\gamma_i(G)$ inductively: $\gamma_1(G) = G$; $\gamma_{i+1}(G) = [\gamma_i(G), G]$.

It is easy to see that $\gamma_{i+1}(G) \subset \gamma_i(G)$. Since $[\gamma_i(G), G] = \gamma_{i+1}(G)$, Exercise 6.36 says that $\gamma_{i+1}(G)$ is a normal subgroup of G for all i. Thus, if $x \in G$ and $y \in \gamma_i(G)$, then x and y commute modulo $\gamma_{i+1}(G)$.

Definition The **descending central series** of G is the normal series

$$G = \gamma_1(G) \supset \gamma_2(G) \supset \cdots.$$

Definition We define a chain of subgroups inductively: $Z^0(G) = \{1\}$; $Z^{i+1}(G)$ is the subgroup of G corresponding to the center of $G/Z^i(G)$:

$$\begin{array}{c} G \to G/Z^i(G) \\ Z^{i+1}(G) \leftrightarrow \text{center} = Z^{i+1}/Z^i \\ Z^i(G) \to \{1\} \end{array}$$

$Z^i(G)$ is the ith **higher center** of G.

By the correspondence theorem, each $Z^i(G) \subset Z^{i+1}(G)$, and $Z^i(G)$ is a normal subgroup of G.

Definition The **ascending central series** of G is

$$\{1\} = Z^0(G) \subset Z^1(G) \subset \cdots.$$

When no confusion can occur, we shall abbreviate $Z^i(G)$ by Z^i and $\gamma_i(G)$ by γ_i.

Theorem 6.20 *For any group G, $Z^m(G) = G$ if and only if $\gamma_{m+1}(G) = \{1\}$. Moreover,*

$$\gamma_{i+1}(G) \subset Z^{m-i}(G) \qquad \textit{for all } i.$$

Proof Assuming $Z^m = G$, we shall prove the inclusion holds by an induction on i. Both terms equal G when $i = 0$, so the induction begins. If $\gamma_{i+1} \subset Z^{m-1}$, then

$$\gamma_{i+2} = [\gamma_{i+1}, G] \subset [Z^{m-i}, G] \subset Z^{m-i-1},$$

the last inclusion following from Exercise 6.38. Since the inclusion holds for all i, it holds for $i = m$. Therefore,

$$\gamma_{m+1} \subset Z^0 = \{1\}.$$

Assuming that $\gamma_{m+1} = \{1\}$, we shall prove by induction on j that $\gamma_{m+1-j} \subset Z^j$ (this is the same inclusion as in the statement of the

theorem, for the sum of the indices is $m + 1$). Both terms equal $\{1\}$ when $j = 0$, so the induction begins. If $\gamma_{m+1-j} \subset Z^j$, then there is a homomorphism

$$G/\gamma_{m+1-j} \to G/Z^j$$

which is onto. Now $\gamma_{m-j}/\gamma_{m+1-j} \subset Z(G/\gamma_{m+1-j})$, since $[\gamma_{m-j}, G] \subset \gamma_{m+1-j}$, and Exercise 6.38 holds. By Exercise 6.39,

$$\gamma_{m-j}Z^j/Z^j \subset Z(G/Z^j) = Z^{j+1}/Z^j.$$

Therefore, $Z^{j+1} \supset \gamma_{m-j}Z^j \supset \gamma_{m-j}$, as desired. Since the inclusion holds for all j, it holds for $j = m$. Therefore,

$$G = \gamma_1 \subset Z^m. \quad \blacksquare$$

Definition A group G is **nilpotent**[3] if there is some integer m such that $\gamma_{m+1}(G) = \{1\}$. The least such integer m is the **class** of G.

A group G is $\{1\}$ if it has class 0; it is abelian and $\neq \{1\}$ if it has class 1. A group G of class 2 is also called **metabelian**[4] and, by Theorem 6.20, is described by $\gamma_2(G) = G' \subset Z(G)$. Thus, Exercise 5.41 says that every nonabelian group of order p^3 is metabelian.

A group G is nilpotent if the descending central series reaches $\{1\}$ or, equivalently, if the ascending central series reaches G. The nilpotent groups form a class of groups lying strictly between the abelian groups and the solvable groups, and should be regarded as generalizations of p-groups.

It is easy to prove by induction that $G^{(i)} \subset \gamma_i(G)$ for all i; it follows that if $\gamma_m(G) = \{1\}$ for some m, then $G^{(m)} = \{1\}$, i.e., if G is nilpotent, then G is solvable. The group S_3 is a solvable group; we claim that it is not nilpotent. Every nilpotent group G has a nontrivial center: if m is the first integer for which $\gamma_{m+1}(G) = \{1\}$, then $\{1\} \neq \gamma_m(G) \subset Z^1(G) = Z(G)$. Therefore, S_3 is not nilpotent, for it is centerless.

Theorem 6.21 *Every subgroup of a nilpotent group G is nilpotent.*

Proof It is easily proved by induction that if $H \subset G$, then $\gamma_i(H) \subset \gamma_i(G)$ for all i. Therefore, $\gamma_m(G) = \{1\}$ implies $\gamma_m(H) = \{1\}$. $\blacksquare$

Theorem 6.22 *If G is nilpotent and $H \lhd G$, then G/H is nilpotent.*

[3] This terminology is borrowed from the theory of Lie algebras.

[4] Many authors define a group G to be **metabelian** if G' is abelian.

Proof If $f: G \to L$ is a homomorphism onto, then it is easily proved by induction that $\gamma_i(L) \subset f(\gamma_i(G))$ for all i. The theorem follows if we take f to be the natural map of G onto G/H. ▮

We have just proved the analogs for nilpotent groups of Theorems 6.11 and 6.12. Is the analog of Theorem 6.13 true? If $H \lhd G$ and both H and G/H are nilpotent, is G nilpotent? Again, S_3 is a counterexample; both A_3 and S_3/A_3 are abelian and hence nilpotent, but we have already seen that S_3 is not nilpotent. The analog of Corollary 6.14, however, is true.

Theorem 6.23 *A direct product G of a finite number of nilpotent groups is nilpotent.*

Proof An induction on the number of direct factors allows us to assume that $G = H \times K$. Another induction proves that

$$\gamma_i(H \times K) \subset \gamma_i(H) \times \gamma_i(K) \qquad \text{for all } i.$$

Let $M = \max\{m, n\}$, where $\gamma_m(H) = \{1\} = \gamma_n(K)$. Then $\gamma_M(H \times K) = \{1\}$, so that $H \times K$ is nilpotent. ▮

Lemma 6.24 *Every finite p-group G is nilpotent.*

Proof We know, by Theorem 5.4, that G and all its nontrivial quotients have nontrivial centers. Therefore, if $Z^i \neq G$ for some i, then $Z^i \subsetneq Z^{i+1}$. Since G is finite, we cannot have this inequality for all i. It follows that $Z^i = G$ for some i, i.e., G is nilpotent. ▮

This lemma is false without the finiteness assumption, for there exist infinite p-groups that are centerless.

Lemma 6.25 *Let G be nilpotent and let H be a proper subgroup. Then $H \neq N_G(H)$.*

Proof There exists an i such that $H \supset \gamma_{i+1}(G)$ but $H \not\supset \gamma_i(G)$ (this is true for any chain of subgroups starting at G and ending at $\{1\}$). Now

$$[\gamma_i, H] \subset [\gamma_i, G] \subset \gamma_{i+1} \subset H,$$

so that $\gamma_i \subset N_G(H)$, by Exercise 6.36. There is thus some element in $N_G(H)$ not in H. ▮

Theorem 6.26 *If G is nilpotent, then every subgroup H of prime index is normal.*

Proof Since $p = [G : H] = [G : N_G(H)][N_G(H) : H]$, one of the factors is 1. By Lemma 6.25, $N_G(H) \neq H$. Therefore $N_G(H) = G$ and $H \lhd G$. ∎

The following theorem is a very satisfying characterization of finite nilpotent groups that indicates they are not too far from p-groups.

Theorem 6.27 *A finite group G is nilpotent if and only if it is the direct product of its sylow subgroups.*

Proof If G is the direct product of its sylow subgroups, then Theorem 6.23 and Lemma 6.24 show G is nilpotent.

To prove the converse, take a sylow subgroup P of G and let $N = N_G(P)$. In Exercise 5.13, we saw that N is its own normalizer. By Lemma 6.25, N cannot be a proper subgroup of G; thus $N_G(P) = G$ and $P \lhd G$. Since all sylow subgroups of G are normal, it follows from Exercise 5.14 that G is the direct product of its sylow subgroups. ∎

EXERCISES

6.40. If G is metabelian and $a \in G$, then the function $[a, \ ]: G \to G$, defined by $b \to [a, b]$, is a homomorphism.

6.41. The group of triangular matrices in Exercise 5.17 is metabelian.

***6.42.** G is nilpotent if and only if there is a normal series

$$G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$$

in which each $G_i \lhd G$ and such that $G_i/G_{i+1} \subset Z(G/G_{i+1})$ for all i. (Such a series is called a **central series**.)

***6.43.** If $H \subset Z(G)$ and G/H is nilpotent, then G is nilpotent.

6.44. The dihedral group D_n is nilpotent if and only if n is a power of 2.

6.45. Let G be a finite nilpotent group of order n, and suppose $m \mid n$. Prove that G contains a subgroup of order m.

6.46. If H, K are normal nilpotent subgroups of a group G, then HK is also a normal nilpotent subgroup of G.

6.47. A finite abelian p-group has a unique subgroup of order p if and only if it is cyclic.

There are many "commutator identities" that are quite useful even though they are quite simple, e.g., $[x, y]^{-1} = [y, x]$. We will need two slightly more complicated identities.

Lemma 6.28 *Let $x, y \in G$ and suppose $[x, y]$ commutes with both x and y. Then, for all n,*

(i) $[x, y]^n = [x^n, y]$;
(ii) $(xy)^n = [y, x]^{n(n-1)/2} x^n y^n$.

Proof To prove (i) for $n \geq 0$, note that

$$\begin{aligned}[x, y]^n[x, y] &= x^{-1}[x, y]^n y^{-1}xy, \text{ by hypothesis} \\ &= x^{-1}[x^n, y]y^{-1}xy, \text{ inductive hypothesis} \\ &= x^{-1}(x^{-n}y^{-1}x^n y)y^{-1}xy \\ &= [x^{n+1}, y].\end{aligned}$$

We claim $[x^{-1}, y] = [x, y]^{-1}$. By hypothesis, $x[y, x] = [y, x]x$, which implies $xy^{-1}x^{-1}y = y^{-1}x^{-1}yx$, i.e., $[x^{-1}, y] = [y, x]$. But we always have $[y, x] = [x, y]^{-1}$. The rest is easy.

To prove the second identity for $n \geq 0$, observe that, by induction,

$$\begin{aligned}(xy)^n(xy) &= [y, x]^{n(n-1)/2} x^n y^n xy \\ &= [y, x]^{n(n-1)/2} x^{n+1}[x, y^{-n}]y^{n+1}.\end{aligned}$$

Using the first identity,

$$[x, y^{-n}] = [y^{-n}, x]^{-1} = [y, x]^n.$$

Therefore

$$\begin{aligned}(xy)^{n+1} &= [y, x]^{n(n-1)/2}[y, x]^n x^{n+1} y^{n+1} \\ &= [y, x]^{(n+1)n/2} x^{n+1} y^{n+1}.\end{aligned}$$

We let the reader prove the identity for negative n. ∎

Theorem 6.29 *Let G be a p-group with a unique subgroup of order p and more than one cyclic subgroup of index p. Then $G \cong Q$, the quaternions.*

Proof If H is a subgroup of G having index p, then $H \lhd G$, by Theorem 6.26. Thus, if $x \in G$, then $xH \in G/H$, a group of order p. Therefore, $x^p \in H$.

Let $A = \langle a \rangle$ and $B = \langle b \rangle$ be distinct cyclic subgroups of index p, and let $D = A \cap B$; $D \lhd G$, being the intersection of normal subgroups. We collect facts about D. First of all, our initial remarks show the subset

$$G^p = \{x^p : x \in G\} \subset D.$$

Now A and B being distinct maximal normal subgroups implies $G = AB$. The product formula gives

$$[G : D] = p^2.$$

It follows that G/D is abelian, so that

$$G' \subset D,$$

by Exercise 2.50. Further, $G = AB$ means that each $x \in G$ is a product of powers of a and b. Thus, an element $d \in D$ is simultaneously a power of a and a power of b, so d commutes with x. Therefore

$$D \subset Z(G).$$

(Indeed, $D = Z(G)$, for $p^2 = [G : D] = [G : Z(G)][Z(G) : D]$, and $[G : Z(G)] \neq p$, by Exercise 3.22.)

Now $G' \subset D \subset Z(G)$, so the hypotheses of Lemma 6.28 hold. Hence, for every $x, y \in G$

$$[y, x]^p = [y^p, x] = 1,$$

because $y^p \in D \subset Z(G)$, and so

$$(xy)^p = [y, x]^{p(p-1)/2} x^p y^p.$$

If p is odd, we have $(xy)^p = x^p y^p$, i.e., $x \to x^p$ is a homomorphism. Setting $G_p = \{x \in G : x^p = 1\}$, Exercise 2.57 says G_p and G^p (defined above) are subgroups and $[G : G_p] = |G^p|$. Thus

$$|G_p| = [G : G^p] = [G : D][D : G^p] \geq p^2.$$

But now G_p, being an elementary abelian subgroup with at least p^2 elements, has more than one subgroup of order p, a contradiction. Therefore, $p = 2$.

Now $G/D = G/Z(G)$ is the 4-group (it cannot be cyclic) with generators aD and bD. Of course, we may replace a and b by other generators of $A = \langle a \rangle$ and $B = \langle b \rangle$, respectively. In particular, $ab^{-m} \notin D$ whenever m is odd. Also, since $D = \langle a \rangle \cap \langle b \rangle$ is cyclic and of index 2 in A and in B, we have $\langle a^2 \rangle = D = \langle b^2 \rangle$, so that $a^2 = b^{2m}$ for some odd m. Finally, note that $D = \langle (ab^{-m})^2 \rangle$, for $\langle ab^{-m} \rangle$ is another subgroup of G of index 2.

As in the first part of the proof, $[y, x]^2 = 1$ for all $x, y \in G$ so that

$$(xy)^4 = [y, x]^6 x^4 y^4 = x^4 y^4.$$

In particular,

$$(ab^{-m})^4 = a^4 b^{-4m} = 1.$$

Since $(ab^{-m})^2$ generates D, we have proved $|D| = 2$, whence $|G| = 8$. To complete the proof, one need only look at the list of groups of order 8. One easily eliminates the abelian groups and D_4, and only Q survives. ▮

Before stating the next theorem, we do a little exercise in congruences.

Lemma 6.30 *Let H_m be the multiplicative group*

$$H_m = \{\bar{a} \in Z_{2^m} : a \text{ is odd}\}.$$

If $m \geq 3$, then

$$H_m = \langle -\bar{1}, \bar{5} \rangle \cong \sigma(2) \times \sigma(2^{m-2}).$$

Proof First note that, by Exercise 2.22, $|H_m| = \varphi(2^m) = 2^{m-1}$. Induction shows that

$$5^{2^{m-3}} = (1 + 4)^{2^{m-3}} \equiv 1 + 2^{m-1} (\text{mod } 2^m).$$

Since H_m is a 2-group, $\bar{5}$ has order 2^s, where $s \geq m - 2$ (for $1 + 2^{m-1} \not\equiv 1 \pmod{2^m}$).

Now $-\bar{1}$ has order 2. We claim $\langle -\bar{1} \rangle \cap \langle \bar{5} \rangle = \{1\}$. For any v, $5^v \equiv 1 \pmod 4$. If $\bar{5}^v \in \langle -\bar{1} \rangle \cap \langle \bar{5} \rangle$, then $5^v \equiv -1 \pmod 4$; but $-1 \equiv 3 \pmod 4$. It follows that these two cyclic subgroups generate their direct product, which is a subgroup of order $2 \cdot 2^s \geq 2 \cdot 2^{m-2} = 2^{m-1} = \varphi(2^m)$. This subgroup is thus all of H_m. ▮

Consider the following group-theoretic situation: G is a group containing elements α and β such that α has order 2^m, $\beta^2 = \alpha^w$, $\beta^{-1}\alpha\beta = \alpha^t$. What can be said about t and w?

First of all,

$$\alpha^w = \beta^2 = \beta^{-1}(\beta^2)\beta = \beta^{-1}(\alpha^w)\beta = \alpha^{tw}.$$

Therefore,

$$w \equiv tw (\text{mod } 2^m).$$

Second, since β^2 commutes with α, we have

$$\alpha = \beta^{-2}\alpha\beta^2 = \beta^{-1}(\alpha^t)\beta = \alpha^{t^2},$$

so that $t^2 \equiv 1 \pmod{2^m}$. Thus, $\bar{t}$ is an element of order 2 in H_m. If $m \geq 3$, Lemma 6.30 says that there are four such elements:

$$t = \pm 1 \quad \text{and} \quad t = \pm 1 + 2^{m-1}.$$

Note that if α and β do not commute, $t = 1$ is ruled out.

Theorem 6.31 *A finite p-group G having a unique subgroup of order p is either cyclic or generalized quaternion.*

Proof The proof is by induction on n, where $|G| = p^n$; of course the theorem is true when $n = 0$.

Assume first that p is odd. If $n > 0$, then G has a subgroup H of index p, by Exercise 5.2, which must be cyclic, by induction. There can be no other subgroup of index p lest it, too, be cyclic and G be the quaternions. Therefore, H is the unique maximal subgroup of G and so it contains every proper subgroup of G. Suppose G is not cyclic. Then, for every $x \in G$, $\langle x \rangle$ is a proper subgroup. Hence $\langle x \rangle \subset H$ and $G \subset H$, a contradiction.

Assume that $p = 2$. We wish to show that G contains a cyclic subgroup of index 2, but first let us show it contains a normal cyclic subgroup of index 4. If $|G| = 16$, then G contains a normal subgroup N of index 4 (Exercise 5.2), and thus of order 4; N is cyclic, for the 4-group has 3 elements of order 2. If $|G| = 2^n > 16$, then a subgroup H of G having index 2 is either cyclic or generalized quaternion, by induction. In either case, H contains a unique (Theorem 6.29) cyclic subgroup $\langle a \rangle$ of order 2^{n-2}. Therefore $\langle a \rangle$ is a characteristic subgroup of H, and hence is a normal cyclic subgroup of G of index 4.

We may assume $|G| = 2^n$ where $n \geq 4$. Now $G/\langle a \rangle$ is a group of order 4, so there exist elements $b, c \in G$ with $b \notin \langle a \rangle$, $c \notin \langle a \rangle$, $b^{-1}c \notin \langle a \rangle$, and with $\langle a, b \rangle$ and $\langle a, c \rangle$ having order 2^{n-1}. If one of $\langle a, b \rangle$, $\langle a, c \rangle$ is cyclic, good; otherwise, both are generalized quaternion, by induction. We may assume there are equations

$$a^{2^{n-2}} = 1, \qquad b^{-1}ab = a^{-1} = c^{-1}ac, \qquad b^2 = a^{2^{n-3}} = c^2.$$

The middle equation $b^{-1}ab = c^{-1}ac$ shows that $b^{-1}c \in C_G(a)$; since $b^{-1}c \notin \langle a \rangle$, the subgroup $\langle a, b^{-1}c \rangle$ is abelian, of order at least 2^{n-1}, and hence cyclic.

We have shown that G must contain a cyclic subgroup $\langle \alpha \rangle$ of index 2. Choose $\beta \in G$, $\beta \notin \langle \alpha \rangle$. Since $[G:\langle \alpha \rangle] = 2$, we have $\beta^2 \in \langle \alpha \rangle$. Changing generator of $\langle \alpha \rangle$ if necessary, we have

$$\beta^2 = \alpha^{2^r}.$$

Note that $r \leq n - 2$ (if $r > n - 2$, then β is a second element of order 2). Furthermore, we may suppose α and β do not commute, otherwise G is abelian, and hence cyclic. The little exercise in congruences gives

$$\beta^{-1}\alpha\beta = \alpha^t,$$

where either $t = -1$, $-1 + 2^{n-2}$, or $1 + 2^{n-2}$ (since α has order 2^{n-1}). We eliminate the last two possible values of t.

Suppose $t = 1 + 2^{n-2}$. For any integer m,

$$(\alpha^m\beta)^2 = \alpha^m\beta^2(\beta^{-1}\alpha^m\beta) = \alpha^{2s},$$

where $s = 2^{r-1} + m(1 + 2^{n-3})$. Since $1 + 2^{n-3}$ is odd ($n \geq 4$), we may solve the congruence

$$2^{r-1} + m(1 + 2^{n-3}) \equiv 0 \bmod 2^{n-2}.$$

For this choice of m, we have $(\alpha^m\beta)^2 = \alpha^{2s} = \alpha^{2^{n-1}} = 1$. Thus, $\alpha^m\beta \notin \langle\alpha\rangle$ is a new element of order 2.

Suppose $t = -1 + 2^{n-2}$. The little exercise says

$$2^r \equiv 2^r(-1 + 2^{n-2}) (\bmod\ 2^{n-1})$$

so that

$$2^r \equiv 0\ (\bmod\ 2^{n-2})$$

and $r = n - 2$. But now

$$(\alpha\beta)^2 = \alpha\beta^2(\beta^{-1}\alpha\beta) = \alpha^{1+t+2^r}$$
$$= \alpha^{2^{n-2}}\alpha^{2^{n-2}} = 1,$$

so a second element of order 2 has been exhibited.

Therefore, $G = \langle\alpha, \beta\rangle$, where

$$\alpha^{2^{n-1}} = 1, \qquad \beta^{-1}\alpha\beta = \alpha^{-1}, \qquad \beta^2 = \alpha^{2^r}.$$

We need only show $r = n - 2$ to complete the proof. This follows from the little exercise: since $t = -1$, we have $2^r \equiv -2^r\ (\bmod\ ^{2^{n-1}})$ so that $2^{r+1} \equiv 0\ (\bmod\ 2^{n-1})$, and $r = n - 2$. ∎

EXERCISES

6.48. Every subgroup of Q_n is either cyclic or generalized quaternion.

6.49. **(Wielandt)** A finite group G is nilpotent if and only if every maximal subgroup is normal. (HINT: Use Exercise 5.13.) Conclude that in this case every maximal subgroup has prime index.

6.50. Let $\Phi(G)$ be the intersection of all the maximal subgroups of G ($\Phi(G)$ is called the **Frattini subgroup** of G). Prove that $\Phi(G) \lhd G$ (indeed, $\Phi(G)$ is even a characteristic subgroup of G).

6.51. If G is finite, prove that $\Phi(G)$ is nilpotent.

6.52. An element $x \in G$ is called a **nongenerator** if, whenever a set $Y \cup \{x\}$ generates a subgroup H, then already Y generates H. Prove that $\Phi(G)$ is the set of all nongenerators of G.

6.53. If G is a finite p-group, then $\Phi(G) = G'G^p$, where G' is the commutator subgroup of G and $G^p = \langle x^p : x \in G \rangle$. Conclude that $G/\Phi(G)$ is a vector space over Z_p.

6.54. **(Burnside Basis Theorem)** If G is a finite p-group, any two minimal generating sets of G have the same number of elements. Moreover, if $x \notin \Phi(G)$, then $\{x\}$ can be extended to a minimal set of generators of G. (HINT: Elements $x_1, \cdots, x_n$ comprise a minimal generating set of G if and only if their cosets modulo $\Phi(G)$ comprise a basis of $G/\Phi(G)$.)

The hypothesis that G is a p-group is necessary. If $G = \langle x \rangle \cong \sigma(6)$, then two minimal generating sets of G are $\{x\}$ and $\{x^2, x^3\}$.

The reader familiar with ring theory will note the analogy of the Frattini subgroup with the Jacobson radical. He will also note the similarity of the Burnside basis theorem with the notion of the rank of a noetherian local ring.

For further results on solvable and nilpotent groups, we refer the reader to the books of Burnside, Hall, Huppert, Kurosh, and Zassenhaus (see Bibliography).

CHAPTER 7

Extensions

THE EXTENSION PROBLEM

If G is a group having a normal subgroup K, then we can "factor" G into the two groups K and G/K. The study of extensions involves the inverse question: Given K and G/K, to what extent may one recapture G?

Definition If K and Q are groups, an **extension** of K by Q is a group G such that:

(i) G contains K as a normal subgroup;
(ii) $G/K \cong Q$.

(As a mnemonic, K denotes kernel and Q denotes quotient.)

In a heuristic sense, an extension G is a "product" of K and Q.

Example 1 Both $\sigma(6)$ and S_3 are extensions of $\sigma(3)$ by $\sigma(2)$.

Example 2 For any two groups K and Q, $K \times Q$ is an extension of K by Q (and also of Q by K).

Example 3 Every extension of a solvable group by a solvable group is itself solvable; an extension of a nilpotent group by a nilpotent group need not be nilpotent (but see Exercise 6.43).

The extension problem formulated by Hölder is, given K and Q, to determine all extensions of K by Q. We can better understand the Jordan-Hölder theorem in the light of this problem. Let G be a group with composition series

$$G = K_0 \supset K_1 \supset \cdots \supset K_{n-2} \supset K_{n-1} \supset K_n = \{1\}$$

and corresponding factor groups

$$K_0/K_1 = Q_1, \cdots, K_{n-1}/K_n = Q_n.$$

Now $K_{n-1} = Q_n$, since $K_n = \{1\}$, but something more interesting happens at the next stage: $K_{n-2}/K_{n-1} = Q_{n-1}$, so that K_{n-2} is an extension of K_{n-1} by Q_{n-1}. If we could solve the extension problem, we could recapture K_{n-2} from K_{n-1} and Q_{n-1}, i.e., from Q_n and Q_{n-1}. Once we have K_{n-2}, we can attack K_{n-3} in a similar manner, for $K_{n-3}/K_{n-2} = Q_{n-2}$. Thus, a solution of the extension problem allows us to recapture K_{n-3} from Q_n, Q_{n-1}, and Q_{n-2}. We continue climbing up the composition series until we reach $K_0 = G$; to do this, we need extensions and the factor groups $Q_1, \cdots, Q_n$. The group G is thus a "product" of the Q_i, and the Jordan-Hölder theorem says that the simple groups occurring as factors in this "factorization" of G are uniquely determined by G. We could thus survey all finite groups if we knew all finite simple groups and if we could solve the extension problem. In particular, we could survey all finite solvable groups if we could solve the extension problem.

A solution of the extension problem consists of determining all groups G with $G/K \cong Q$. But what does "determining" a group G mean? We gave two answers to this question in Chapter 1 when we considered "knowing" a group G. One answer is that a multiplication table for G can be constructed; a second answer is that the isomorphism class of G can be characterized. Schreier solved the extension problem in the first sense; given K and Q, all multiplication tables of extensions G of K by Q can be constructed. On the other hand, no solution is known in the second sense. For example, given K and Q, Schreier's solution does not allow us to calculate the number of nonisomorphic extensions of K by Q.

We shall see that, in essence, the extensions of K by Q themselves form an abelian group! The computation of one of these groups is the Schur-Zassenhaus lemma (Theorem 7.19). More general techniques for computing groups of extensions will be considered in Chapter 10, where we discuss homological algebra.

EXERCISES

7.1. Any two extensions of K by Q have the same number of elements, namely, $|K||Q|$.

7.2. There are exactly two nonisomorphic extensions of $\sigma(3)$ by $\sigma(2)$; there is exactly one extension of $\sigma(2)$ by $\sigma(3)$.

7.3. Every nonabelian group of order p^3, p prime, is an extension of $\sigma(p)$ by $\sigma(p) \times \sigma(p)$.

7.4. If G is an extension of K by Q, must G contain a subgroup isomorphic to Q?

7.5. If $(a, b) = 1$ and if K and Q are abelian groups of orders a and b, respectively, prove that there is only one abelian extension of K by Q.

AUTOMORPHISM GROUPS

The coming construction is essential for a discussion of extensions; it is also of great intrinsic interest.

Definition The **automorphism group** of a group G, denoted $\mathrm{Aut}(G)$, is the set of all automorphisms of G under the binary operation of composition.

It is easy to check that $\mathrm{Aut}(G)$ is a group; indeed, $\mathrm{Aut}(G)$ is a subgroup of S_G.

Definition An automorphism α of G is **inner** if it is conjugation by an element of G, i.e., $\alpha(x) = f_a(x) = axa^{-1}$ for some $a \in G$; otherwise, α is **outer**.

Theorem 7.1 *The set of all inner automorphisms of G, denoted* $\mathrm{Inn}(G)$, *is a normal subgroup of* $\mathrm{Aut}(G)$; *furthermore,*

$$G/Z(G) \cong \mathrm{Inn}(G).$$

Proof If $a \in G$, let f_a denote conjugation by a. The function $a \to f_a$ is easily seen to be a homomorphism of G into $\mathrm{Aut}(G)$. The image of this map is clearly $\mathrm{Inn}(G)$, so that $\mathrm{Inn}(G)$ is a subgroup of $\mathrm{Aut}(G)$. The kernel of this map is $Z(G)$, for the following statements are equivalent: a is in the kernel; f_a is the identity function on G; $axa^{-1} = x$ for all $x \in G$; $a \in Z(G)$. It follows from the first isomorphism theorem that $G/Z(G) \cong \mathrm{Inn}(G)$.

In order to prove that $\mathrm{Inn}(G) \lhd \mathrm{Aut}(G)$, we must show that if $\alpha \in \mathrm{Aut}(G)$, then $\alpha f_a \alpha^{-1}$ is an inner automorphism. In fact, $\alpha f_a \alpha^{-1} = f_{\alpha(a)}$, as the reader should check. ∎

EXERCISES

*7.6. Prove that $\text{Aut}(S_3) \cong S_3$.

7.7. Prove that $\text{Aut}(V) \cong S_3$, where V is the 4-group. Conclude that nonisomorphic groups can have isomorphic automorphism groups.

*7.8. Let G be an elementary abelian group of order p^n, where p is prime. Prove that $\text{Aut}(G) \cong GL(n, p)$.

7.9. Let H and K be finite groups whose orders are relatively prime. Prove that $\text{Aut}(H \times K) \cong \text{Aut}(H) \times \text{Aut}(K)$.

*7.10. Let G be a finite group with $\text{Aut}(G) = \{1\}$. Prove that G has at most two elements. (The finiteness condition is unnecessary, but the reader may not possess the tools for dealing with the general case until Chapter 9.)

7.11. Let H be a subgroup of G. Then $C_G(H) \lhd N_G(H)$ and $N_G(H)/C_G(H)$ can be imbedded in $\text{Aut}(H)$. (Theorem 7.1 is the special case of this when $H = G$.)

7.12. If p is prime, then $\text{Aut}(\sigma(p)) \cong \sigma(p - 1)$. (HINT: Use Exercise 4.17.)

Definition A **unit** in a commutative ring with unit R is an element having a multiplicative inverse in R. The **group of units**, $U(R)$, is the multiplicative group of all units in R.

EXERCISES

**7.13. Prove that $\text{Aut}(\sigma(n)) \cong U(Z_n)$ and hence has order $\varphi(n)$. (HINT: An integer t represents a unit in Z_n if and only if $(t, n) = 1$.)

**7.14. If p is prime, then $a \equiv b \pmod{p}$ implies

$$a^{p^{n-1}} \equiv b^{p^{n-1}} \pmod{p^n}.$$

Theorem 7.2

$$\text{Aut}(\sigma(2^m)) \cong \begin{cases} \{1\} & \text{if } m = 1 \\ \sigma(2) & \text{if } m = 2 \\ \sigma(2) \times \sigma(2^{m-2}) & \text{if } m \geq 3. \end{cases}$$

If p is an odd prime,

$$\text{Aut}(\sigma(p^n)) \cong \sigma((p - 1)p^{n-1}).$$

Proof By Exercise 7.13, $\mathrm{Aut}(\sigma(2^m)) \cong U(Z_{2^m})$ and has order $\varphi(2^m)$. Since $\varphi(2) = 1$ and $\varphi(4) = 2$, the theorem holds when $m \leq 2$. If $m \geq 3$, then this theorem is precisely Lemma 6.30.

Assume that p is odd and that $n \geq 2$ (we already know the theorem for $n = 1$). Denote $U(Z_{p^n})$ by G. Now $|G| = \varphi(p^n) = (p-1)p^{n-1}$; we shall show that $G \cong \sigma(p-1) \times \sigma(p^{n-1})$.

Let $B = \{\bar{b} \in G: b \equiv 1 \pmod{p}\}$; note that B is a subgroup of G. If $1 \leq b < p^n$, then b has a unique expression in base p:

$$b = b_0 + b_1 p + \cdots + b_{n-1}p^{n-1}, \qquad 0 \leq b_i \leq p-1.$$

It follows that $|B| = p^{n-1}$, for $b \in B$ if and only if $b_0 = 1$. We claim $B = \langle \overline{1+p} \rangle$. Since $|B| = p^{n-1}$, it suffices to show $(1+p)^{p^{n-2}} \not\equiv 1 \pmod{p^n}$, for then $\overline{1+p}$ must have order p^{n-1}. The incongruence is proved by induction on n, the case $n = 2$ being trivial.

Since $1 + p \equiv 1 \pmod{p}$, Exercise 7.14 gives

$$(1+p)^{p^{n-3}} \equiv 1 \pmod{p^{n-2}},$$

so that, by induction,

$$(1+p)^{p^{n-3}} = 1 + kp^{n-2} \qquad \text{where } k \not\equiv 0 \pmod{p}.$$

Therefore

$$\begin{aligned}(1+p)^{p^{n-2}} &= (1+kp^{n-2})^p \\ &= 1 + \binom{p}{1} kp^{n-2} + \cdots + k^p p^{p(n-2)} \\ &\equiv 1 + kp^{n-1} \pmod{p^n} \\ &\not\equiv 1 \pmod{p^n}.\end{aligned}$$

By the primary decomposition, $G = A \times B$, where A is a subgroup of order $p-1$. Consider the map $f: G \to U(Z_p)$ defined by $f(\bar{a}) = \mathrm{cls}\, a$, where cls a means the congruence class of a modulo p. Clearly f is onto with kernel B, so that $G/B \cong U(Z_p) \cong \sigma(p-1)$. Since $G/B \cong A$, we have $A \cong \sigma(p-1)$. Therefore $G \cong \sigma(p-1) \times \sigma(p^{n-1}) \cong \sigma((p-1)p^{n-1})$, by Exercise 4.7. ∎

EXERCISES

7.15. Give a complete set of invariants describing $U(Z_n)$.

7.16. Let X be the set of all conjugacy classes of a group G. Every automorphism α of G induces a permutation of X. Conclude that $\mathrm{Aut}(G)$ is a permutation group on X as well as a permutation

group on G. This remark is the origin of the notion of **imprimitive** permutation groups (see Burnside, Passman, and Wielandt).

**7.17. Let G be a finite group and let C be a conjugacy class of G consisting of h elements of order t. If no other conjugacy class of G comprised of elements of order t has exactly h elements, then $\alpha(C) = C$ for every $\alpha \in \mathrm{Aut}(G)$.

Theorem 7.1 suggests the following class of groups.

Definition A group G is **complete** in case $Z(G) = \{1\}$ and every automorphism of G is inner.

If G is a complete group, then $\mathrm{Aut}(G) \cong G$. An interesting generalization of Exercise 7.6 is that almost all the symmetric groups are complete.

Lemma 7.3 *If $\alpha: S_n \to S_n$ is an automorphism that preserves transpositions, i.e., if α sends any transposition into a transposition, then α is inner.*

Proof Recall that two elements of S_n are conjugate if and only if they have the same cycle structure, so that every conjugation of S_n preserves transpositions.

If $\pi \in S_n$ and $\gamma: S_n \to S_n$, denote $\gamma(\pi)$ by π^γ.

We shall prove by induction on t that there exist conjugations $\beta_2, \cdots, \beta_t$ such that $\beta_t^{-1} \cdots \beta_2^{-1}\alpha$ fixes $(1\ 2), \cdots, (1\ t)$. Now $(1\ 2)^\alpha = (ij)$; define β_2 to be conjugation by $(1\ i)(2j)$ (where our notation sets, e.g., $(1\ i) =$ identity if $1 = i$). Our quick way of computing conjugations (Lemma 3.9) shows that $(1\ 2)^\alpha = (1\ 2)^{\beta_2}$, and so $\beta_2^{-1}\alpha$ fixes $(1\ 2)$.

Let $\gamma = \beta_t^{-1} \cdots \beta_2^{-1}\alpha$ be given by the inductive hypothesis. Since γ preserves transpositions, $(1\ t+1)^\gamma = (l\ k)$. Now $\{1, 2\} \cap \{l, k\} \neq \varnothing$, otherwise $[(1\ 2)(1\ t+1)]^\gamma = (1\ 2)(l\ k)$ has order 2 while $(1\ 2)(1\ t+1)$ has order 3. Thus $(1\ t+1)^\gamma = (1\ k)$ or $(2\ k)$. We must have $k > t$, because γ is one-one. Define $\beta_{t+1}: S_n \to S_n$ to be conjugation by $(k\ t+1)$. Now β_{t+1} fixes $(1\ 2), \cdots, (1\ t)$ and $(1\ t+1)^{\beta_{t+1}} = (1\ t+1)^\gamma$, so that $\beta_{t+1}^{-1}\gamma = \beta_{t+1}^{-1}\beta_t^{-1} \cdots \beta_2^{-1}\alpha$ fixes $(1\ 2), \cdots, (1\ t+1)$.

It follows that $\beta_n^{-1} \cdots \beta_2^{-1}\alpha$ fixes $(1\ 2), \cdots, (1\ n)$. Since these $n-1$ transpositions generate S_n, by Exercise 3.20, $\beta_n^{-1} \cdots \beta_2^{-1}\alpha$ is the identity, and $\alpha = \beta_2 \cdots \beta_n$ is inner. ∎

Theorem 7.4 *S_n is complete if $n \neq 2$ and $n \neq 6$.*

Proof We remark that $S_2 \cong \sigma(2)$ is not complete because it has a center. If $n \geq 3$, we have seen in Exercise 3.34 that S_n is centerless.

Let C_1 be the conjugacy class of S_n consisting of all transpositions; every element in C_1 has order 2. Since every permutation is a product of disjoint cycles, the only elements in S_n of order 2 are products of disjoint transpositions; let C_k be the conjugacy class consisting of all products of k disjoint transpositions. We shall show that if $n \neq 6$, the size of each C_k $(k \neq 1)$ is distinct from the size of C_1. Exercise 7.17 will then give $\alpha(C_1) = C_1$ for every $\alpha \in \mathrm{Aut}(S_n)$, i.e., every $\alpha \in \mathrm{Aut}(S_n)$ preserves transpositions. An application of Lemma 7.3 will then complete the proof.

There are $n(n-1)/2$ transpositions in S_n. How many *sets* of k disjoint transpositions are there? Answer:

$$\frac{n(n-1)}{2} \cdot \frac{(n-2)(n-3)}{2} \cdots \frac{(n-2k+2)(n-2k+1)}{2}.$$

How many *products* of k disjoint transpositions are there? If $\tau_1, \tau_2, \cdots, \tau_k$ are disjoint transpositions, they commute, and so the order in which they are written is irrelevant. There being $k!$ different orderings, the size of C_k is

$$\frac{1}{k!}\frac{1}{2^k} n(n-1)(n-2)\cdots(n-2k+1).$$

The problem is now reduced to the question: Can the size of C_k equal $n(n-1)/2$? Equivalently, does the following equation hold for $k > 1$?

$$(*) \qquad (n-2)(n-3)\cdots(n-2k+1) = k!\,2^{k-1}.$$

Since the left side is positive, we must have $n \geq 2k$. Therefore, for fixed n,

$$\text{left side} \geq (2k-2)(2k-3)\cdots(2k-2k+1) = (2k-2)!.$$

An easy induction shows that if $k \geq 4$, then

$$(2k-2)! > k!2^{k-1},$$

so that (*) can hold only if $k < 4$, regardless of the value of n. We may now assume that $k = 2$ or $k = 3$. It is easy to see that (*) never holds if $k = 2$, so that only the case $k = 3$ remains. Now, since $n \geq 2k$, we must have $n \geq 6$. If $n > 6$, the left side of (*) $> 5 \cdot 4 \cdot 3 \cdot 2 = 120$, while the right side of (*) $= 3!2^2 = 24$. (If

$n = 6$ and $k = 3$, then (*) does hold.) We have shown that if $n \neq 6$, then there is no $k > 1$ for which (*) holds. Hence, if $n \neq 6$, the size of each C_k ($k \neq 1$) is distinct from the size of C_1. ∎

Corollary 7.5 $\mathrm{Aut}(S_n) \cong S_n$ *for* $n \neq 2$ *and* $n \neq 6$.

It is known that $\mathrm{Aut}(S_6)/\mathrm{Inn}(S_6) \cong \sigma(2)$, so that there is essentially only one outer automorphism of S_6. An explicit construction of an outer automorphism can be found in D. W. Miller, "On a Theorem of Hölder," *Amer. Math. Monthly*, 65 (1958), pp. 252–254.

The next theorem provides another source of (possibly infinite) complete groups.

Theorem 7.6 *If G is a simple nonabelian group, then* $\mathrm{Aut}(G)$ *is complete.*

Proof Let $I = \mathrm{Inn}(G) \lhd \mathrm{Aut}(G) = A$. Since G is simple and not abelian, $Z(G) = \{1\}$, and so $G \cong I$. First of all, $Z(A) = \{1\}$. In fact, we show $C_A(I) = \{1\}$: if $\alpha \in C_A(I)$, then $f_b = \alpha f_b \alpha^{-1} = f_{\alpha(b)}$, all $b \in G$; hence $\alpha(b)b^{-1} \in Z(G) = \{1\}$, all $b \in G$, and $\alpha = 1$.

To see that every automorphism of A is inner, suppose $\sigma \in \mathrm{Aut}(A)$. Since $I \lhd A$, we have $\sigma(I) \lhd A$, so that $I \cap \sigma(I) \lhd \sigma(I)$. Simplicity of $\sigma(I)$ gives $I \cap \sigma(I) = \{1\}$ or $I \cap \sigma(I) = \sigma(I)$. In the first case,

$$[I, \sigma(I)] \subset I \cap \sigma(I) = \{1\},$$

which implies $\sigma(I) = \{1\}$, by paragraph one. This contradiction shows that $I \cap \sigma(I) = \sigma(I)$, i.e., $\sigma(I) \subset I$. Now, if $f_b \in I$, then $\sigma(f_b) \in I$; therefore,

$$\sigma(f_b) = f_{\alpha(b)},$$

where, as one easily checks, $\alpha \in \mathrm{Aut}(G) = A$.

Let $\tau = \sigma \circ (F_\alpha)^{-1}$, where $F_\alpha : A \to A$ is conjugation by α. Observe that

$$\begin{aligned} \tau(f_b) &= \sigma F_\alpha^{-1}(f_b) = \sigma(\alpha^{-1} f_b \alpha) \\ &= \sigma(f_{\alpha^{-1}(b)}) \\ &= f_{\alpha\alpha^{-1}(b)} = f_b. \end{aligned}$$

Thus, τ fixes everything in I.

We claim that τ is the identity on A. If not, $\tau(\beta) \neq \beta$ for some $\beta \in A$. For every $b \in G$,

$$\begin{aligned} \beta f_b \beta^{-1} &= \tau(\beta f_b \beta^{-1}) && \text{(since } I \lhd A \text{ and } \tau \text{ fixes } I\text{)} \\ &= \tau(\beta) f_b \tau(\beta)^{-1} && \text{(since } \tau \text{ fixes } I\text{).} \end{aligned}$$

Hence $\tau(\beta)\beta^{-1} \in C_A(I) = \{1\}$ and $\tau(\beta) = \beta$, a contradiction. Therefore $\tau = 1$ and σ is the inner automorphism F_α. ∎

It follows that $\mathrm{Aut}(\mathrm{Aut}(G)) \cong \mathrm{Aut}(G)$ whenever G is a nonabelian simple group. There is a beautiful theorem of Wielandt with a similar conclusion. For a centerless group G, we know that G can be imbedded in $\mathrm{Aut}(G)$; moreover, $\mathrm{Aut}(G)$ is also centerless. Thus, this process may be iterated,

$$G \subset \mathrm{Aut}(G) \subset \mathrm{Aut}(\mathrm{Aut}(G)) \subset \cdots,$$

to give the **automorphism tower** of G. Wielandt proved that if G is a finite centerless group, then its automorphism tower is constant from some point on.

Theorem 7.7 *Let $H \lhd G$, where H is a complete group. Then H is a direct factor of G, i.e., there exists a normal subgroup K of G with $G = H \times K$.*

Proof If $g \in G$, then $g^{-1}hg \in H$ for all $h \in H$, since $H \lhd G$. The function $h \to g^{-1}hg$ is thus an automorphism of H. Since every automorphism of H is inner, there exists an element $\eta = \eta(g) \in H$ with $\eta^{-1}h\eta = g^{-1}hg$ for all $h \in H$. Let $K = C_G(H)$, the centralizer of H in G. Clearly K contains each $g\eta^{-1}$. We claim that $G = H \times K$. First of all, $HK = G$, for if $g \in G$, then $g = (g\eta^{-1})\eta$. Second, if $x \in H \cap K$, then $x \in Z(H) = \{1\}$, since H is centerless. Now $g^{-1}Kg = \eta^{-1}(\eta g^{-1})K(g\eta^{-1})\eta \subset \eta^{-1}K\eta$, since $g\eta^{-1} \in K$; furthermore, $\eta^{-1}K\eta \subset K$, by the definition of K. It follows that $K \lhd G$, and hence $G = H \times K$. ∎

EXERCISE

7.18. If G is a complete group and $G' \neq G$, then G is not the commutator subgroup of any group containing it. Conclude that S_n, $n \neq 2, 6$, is never a commutator subgroup.

SEMIDIRECT PRODUCTS

A group G is the direct product of two normal subgroups K and Q in case $K \cap Q = \{1\}$ and $KQ = G$. A natural generalization of direct products is the situation in which only one of the subgroups is required to be normal.

Definition A group G is a **semidirect product** of K by Q in case G contains subgroups K and Q such that:

$$\begin{aligned} &\text{(i)} \quad K \lhd G; \\ &\text{(ii)} \quad KQ = G; \\ &\text{(iii)} \quad K \cap Q = \{1\}. \end{aligned}$$

It follows from the second isomorphism theorem that a semidirect product of K by Q is an extension of K by Q; a semidirect product is often called a **split extension**.

EXERCISES

7.19. S_n is a semidirect product of A_n by $\sigma(2)$.

7.20. Let G be a finite solvable group of order ab, where $(a, b) = 1$. If G contains a normal subgroup K of order a, then G is a semidirect product of K by some group Q of order b. (HINT: Use P. Hall's theorem, Theorem 6.19.)

***7.21.** D_n is a semidirect product of $\sigma(n)$ by $\sigma(2)$.

7.22. If p is a prime, then $\sigma(p^n)$ is not a semidirect product.

7.23. The quaternions Q is not a semidirect product.

***7.24.** Both S_3 and $\sigma(6)$ are semidirect products of $\sigma(3)$ by $\sigma(2)$.

Exercise 7.24 is a bit jarring at first, for it says that, in contrast to direct product, a semidirect product of K by Q is not determined up to isomorphism by the two subgroups. When we reflect on this, however, we see that a semidirect product should depend on how K is normal in G.

Lemma 7.8 *Let G be a semidirect product of K by Q. There is a homomorphism $\theta: Q \to \mathrm{Aut}(K)$ defined by*

$$\theta_x(k) = xkx^{-1}, \qquad \text{all } k \in K,\ x \in Q.$$

Moreover,

$$\left.\begin{aligned} \theta_x(\theta_y(k)) &= \theta_{xy}(k) \\ \theta_1(k) &= k \end{aligned}\right\} \qquad k \in K,\ x, y, 1 \in Q.$$

Proof Straightforward, using the normality of K. ∎

The object of our study is to recapture a semidirect product G from K and Q, so let us begin with only K, Q, and some homomorphism $\theta: Q \to \mathrm{Aut}(K)$.

Definition Given K, Q, and $\theta: Q \to \mathrm{Aut}(K)$, then a semidirect product G of K by Q **realizes** θ in case

$$\theta_x(k) = xkx^{-1} \qquad \text{for all } k \in K.$$

In this language, Lemma 7.8 states that every semidirect product G of K by Q determines some θ that G realizes. The intuitive meaning of θ is that it describes how K is normal in G.

Definition Let K, Q, and $\theta: Q \to \mathrm{Aut}(K)$ be given. Then $K \times_\theta Q$ is the set of all ordered pairs $(k, x) \in K \times Q$ under the binary operation

$$(k, x)(k_1, y) = (k\theta_x(k_1), xy).$$

Theorem 7.9 *Let K, Q, and $\theta: Q \to \mathrm{Aut}(K)$ be given; then $G = K \times_\theta Q$ is a semidirect product of K by Q that realizes θ.*

Proof We first prove that $G = K \times_\theta Q$ is a group. Multiplication is associative:

$$\begin{aligned}[(k, x)(k_1, y)](k_2, z) &= (k\theta_x(k_1), xy)(k_2, z) \\ &= (k\theta_x(k_1)\theta_{xy}(k_2), xyz);\end{aligned}$$

$$\begin{aligned}(k, x)[(k_1, y)(k_2, z)] &= (k, x)(k_1\theta_y(k_2), yz) \\ &= (k\theta_x(k_1\theta_y(k_2)), xyz).\end{aligned}$$

The formulas in Lemma 7.8 show that these are the same.

The identity element is $(1, 1)$, for

$$(1, 1)(k, x) = (1 \cdot \theta_1(k), 1x) = (k, x).$$

The inverse of (k, x) is $((\theta_{x^{-1}}(k))^{-1}, x^{-1})$, for

$$(k, x)((\theta_{x^{-1}}(k))^{-1}, x^{-1}) = (k\theta_x(\theta_{x^{-1}}(k))^{-1}, xx^{-1}) = (1, 1).$$

Let us identify K with the subset of G consisting of all pairs of the form $(k, 1)$. Since the only "twist" occurs in the first coordinate, the map $\pi: G \to Q$ defined by $\pi(k, x) = x$ is a homomorphism. It is quickly checked that $\ker \pi = K$, so that K is a normal subgroup of G.

Identify Q with all pairs $(1, x)$. Then Q is a subgroup of G with $KQ = G$ and $K \cap Q = \{(1, 1)\}$. Therefore, G is a semidirect product of K by Q.

To see that G realizes θ, compute:

$$(1, x)(k, 1)(1, x)^{-1} = (\theta_x(k), x)(1, x^{-1}) = (\theta_x(k), 1). \quad \blacksquare$$

Since $G = K \times_\theta Q$ does realize θ, there can be no confusion if we write k^x instead of $\theta_x(k)$. The binary operation of $K \times_\theta Q$ may be now written:

$$(k, x)(k_1, y) = (kk_1^x, xy).$$

Theorem 7.10 *If G is a semidirect product of K by Q, then $G \cong K \times_\theta Q$ for some $\theta: Q \to \mathrm{Aut}(K)$.*

Proof As in Lemma 7.8, define $\theta_x(k) = xkx^{-1}$. Since $G = KQ$, each $g \in G$ has an expression $g = kx$, where $k \in K$ and $x \in Q$; this expression is unique because $K \cap Q = \{1\}$. Multiplication in G satisfies

$$\begin{aligned}(kx)(k_1x_1) &= k(xk_1x^{-1})xx_1 \\ &= kk_1^x xx_1.\end{aligned}$$

It is now easy to see that the map $K \times_\theta Q \to G$ defined by $(k, x) \to kx$ is an isomorphism. ∎

To illustrate how these results may be used, we construct the group T of order 12 defined in Exercise 5.40 and the nonabelian group of order p^3 in which every element has order p.

Example 4 The group T is a semidirect product of $\sigma(3)$ by $\sigma(4)$.

Let $\sigma(3) = \langle k \rangle$ and let $\sigma(4) = \langle x \rangle$. Define $\theta: \sigma(4) \to \mathrm{Aut}(\sigma(3)) \cong \sigma(2)$ by sending x into the generator. In detail,

$$k^x = k^2 \qquad \text{and} \qquad (k^2)^x = k,$$

while x^2 acts on $\langle k \rangle$ as the identity automorphism.

The group $\sigma(3) \times_\theta \sigma(4)$ has order 12. If $a = (k^2, x^2)$ and $b = (1, x)$, then the reader may verify that

$$a^6 = 1 \qquad \text{and} \qquad b^2 = a^3 = (ab)^2,$$

which are the relations of T.

Example 5 Let p be a prime, $K = \langle k_1, k_2 \rangle$ be elementary abelian of order p^2, and $Q = \langle x \rangle$ be cyclic of order p. Define $\theta: Q \to \mathrm{Aut}(K) \cong GL(2, p)$ by

$$x^i \to \begin{bmatrix} 1 & 0 \\ i & 1 \end{bmatrix}.$$

Thus, $k_1^x = k_1k_2$ and $k_2^x = k_2$. The commutator $k_1^x k_1^{-1}$ is seen to be k_2. Therefore, $G = K \times_\theta Q$ is a group of order p^3, $G = \langle k_1, k_2, x \rangle$, and these generators satisfy relations: $k_1^p = k_2^p = x^p = 1$, $k_2 = xk_1x^{-1}k_1^{-1}$, and $[k_2, k_1] = 1 = [k_2, x]$.

If p is odd, we have the nonabelian group of order p^3 in which every element has order p; if $p = 2$, we have D_4. Note, using Exercise 7.21, that $D_4 \cong \sigma(4) \times_\psi \sigma(2)$ and $D_4 \cong V \times_\theta \sigma(2)$, where V is the 4-group; a group may thus have distinct factorizations as a semidirect product.

EXERCISES

7.25. $K \times_\theta G \cong K \times Q$, the direct product, if and only if θ is trivial ($\theta_x = 1_K$ for all $x \in Q$).

7.26. If p and q are distinct primes, construct all semidirect products of $\sigma(p)$ by $\sigma(q)$. Compare your results with Theorem 5.13; the condition $q \mid p - 1$ should now be more understandable.

Definition The **holomorph** of K is $K \times_\theta \mathrm{Aut}(K)$, where $\theta\colon \mathrm{Aut}(K) \to \mathrm{Aut}(K)$ is the identity map.

EXERCISES

7.27. Let G be the holomorph of K. Prove that every automorphism of K is the restriction of an inner automorphism of G.

7.28. For any group K, a **holomorphism** is a one-one correspondence $h\colon K \to K$ such that

$$h(xy^{-1}z) = h(x)h(y)^{-1}h(z) \qquad \text{for every } x, y, z \in K.$$

(Every automorphism of K is a holomorphism, as is every left or right translation.) Prove that $\mathrm{Hol}(K)$, the set of all holomorphisms of K, is a subgroup of the symmetric group S_K, and that $\mathrm{Hol}(K)$ is isomorphic to the holomorph of K.

****7.29.** Let G be an extension of K by Q and let $\pi\colon G \to Q$ be a map of G onto Q having kernel K. Prove that G is a semidirect product of K by Q if and only if there is a homomorphism $l\colon Q \to G$ with $\pi \circ l = 1_Q$.

Let A and Q be groups with Q finite. Define K to be the direct product of $|Q|$ copies of A. One can think of K either as "vectors" whose coordinates lie in A, or, more formally, as A^Q, all functions $Q \to A$ under pointwise multiplication:

$$(f_1 f_2)(x) = f_1(x)f_2(x) \qquad x \in Q.$$

Define $\theta\colon Q \to \mathrm{Aut}(K)$ by $\theta_y(f) = f^y \in A^Q$, where

$$f^y(x) = f(xy) \qquad x, y \in Q.$$

If f is regarded as a vector with coordinates $f(x)$, $x \in Q$, then f^y is obtained from f by permuting its coordinates via translation by y.

For example, if $Q = \{1, x, y, xy\}$ is the 4-group and $f = (a_1, a_x, a_y, a_{xy}) \in A \times A \times A \times A$, then $f^y = (a_y, a_{xy}, a_1, a_x)$.

Definition Let A and Q be groups, with Q finite.[1] The **wreath product** of A by Q, denoted $A \wr Q$ (or A wr Q), is the semidirect product

$$A \wr Q = A^Q \times_\theta Q,$$

where θ is the homomorphism defined above.

EXERCISES

7.30. If A and Q are finite, then

$$|A \wr Q| = |A|^{|Q|}|Q|.$$

7.31. Wreath product is associative: if Q and R are finite, then

$$(A \wr Q) \wr R \cong A \wr (Q \wr R).$$

7.32. Prove that $D_4 \cong \sigma(2) \wr \sigma(2)$.

7.33. If Q is finite, then $A \wr Q$ contains an isomorphic copy of every extension of A by Q.

Let k and m be positive integers. How many integers $\leq m$ are divisible by k? The answer is t, where $k, 2k, 3k, \cdots, tk \leq m$, while $(t+1)k > m$; thus, $t = [m/k]$, the greatest integer in m/k. If we let p be a prime, what is the largest power of p dividing $m!$? By our initial remarks, $[m/p]$ factors are divisible by p, $[m/p^2]$ factors are divisible by p^2, etc. Therefore, if $m! = p^N m'$, where $(m', p) = 1$, then

$$N = [m/p] + [m/p^2] + [m/p^3] + \cdots$$

(this sum is finite, for $[m/p^k] = 0$ as soon as $p^k > m$). In particular, if $m = p^n$, then $N = p^{n-1} + p^{n-2} + \cdots + p + 1$; we have computed the order of a p-sylow subgroup of the symmetric group S_{p^n}.

Theorem 7.11 *If p is a prime, then a p-sylow subgroup of the symmetric group S_{p^n} is the wreath product of $\sigma(p)$ with itself n times.*

Proof The proof is by induction on n, the case $n = 1$ holding because a p-sylow subgroup of S_p is cyclic of order p. Now suppose $n > 1$. Let B_0 be the set of integers 1 through p^{n-1} and let $B_i = B_0 + ip^{n-1}$. Thus the set of integers 1 through p^n is the disjoint union $B_0 \cup B_1 \cup \cdots \cup B_{p-1}$. Let $\alpha \in S_{p^n}$ be the permutation that "adds p^{n-1} modulo p^n": if $b_0 \in B_0$,

$$\alpha(b_0 + ip^{n-1}) = \begin{cases} b_0 + (i+1)p^{n-1} & \text{if } i < p-1; \\ b_0 & \text{if } i = p-1. \end{cases}$$

[1] If Q is infinite, there are two distinct wreath products, depending on how one defines an infinite product of groups.

Thus, $\alpha(B_0) = B_1$, $\alpha(B_1) = B_2, \cdots, \alpha(B_{p-1}) = B_0$. One easily sees that α has order p.

Let $H = H_0$ be a p-sylow subgroup of $S_{B_0} \cong S_{p^{n-1}}$. Now $\alpha|B_i: B_i \to B_{i+1}$ (subscripts taken modulo p) is a one-one correspondence; hence (Exercise 3.2) it induces an isomorphism $S_{B_i} \cong S_{B_{i+1}}$. Let H_1 be the image of H_0, H_2 the image of H_1, etc. We have thus chosen one p-sylow subgroup for each i.

Now S_{p^n} contains $K = H_0 \times H_1 \times \cdots \times H_{p-1}$ (disjoint permutations commute) and $\langle\alpha\rangle$. Moreover $K \cap \langle\alpha\rangle = \{1\}$, and if $(\beta_0, \beta_1, \cdots, \beta_{p-1}) \in K$, then Lemma 3.9 gives $(\beta_0, \beta_1, \cdots, \beta_{p-1})^\alpha = (\beta_1, \beta_2, \cdots, \beta_0)$. It follows easily that $\langle K, \alpha\rangle \cong H \wr \langle\alpha\rangle \cong H \wr \sigma(p)$. By induction, H is a wreath product of $n - 1$ copies of $\sigma(p)$, so that $\langle K, \alpha\rangle$ is a wreath product of n copies of $\sigma(p)$.

To check that $\langle K, \alpha\rangle \cong H \wr \sigma(p)$ is a p-sylow subgroup, it suffices to see that its order is $p^{N(n)}$, where $N(n) = p^{n-1} + p^{n-2} + \cdots + p + 1$. We know that $|H| = p^{N(n-1)}$; hence $|H \wr \sigma(p)| = (p^{N(n-1)})^p p = p^{pN(n-1)+1} = p^{N(n)}$. ∎

Using Theorem 7.11, one may compute a p-sylow subgroup of any S_n. Write n in base p:

$$n = a_0 + a_1 p + \cdots + a_t p^t \qquad \text{where } 0 \leq a_i \leq p - 1.$$

Now partition the set of integers 1 through n into a_0 singletons, a_1 subsets with p elements, and so forth. On each of these subsets X, construct a p-sylow subgroup of S_X. The direct product of all these subgroups is a p-sylow subgroup of S_n (it is a subgroup because disjoint permutations commute; it is a sylow subgroup because it has the right order). As an illustration, we compute a 2-sylow subgroup of S_6 (this has been done by hand in Exercise 5.20). In base 2, $6 = 0 \cdot 1 + 1 \cdot 2 + 1 \cdot 4$. A 2-sylow subgroup of S_2 is $\sigma(2)$; a 2-sylow subgroup of S_4 is $\sigma(2) \wr \sigma(2) \cong D_4$. Hence, a 2-sylow subgroup of S_6 is isomorphic to $D_4 \times \sigma(2)$.

FACTOR SETS

Since there are nonsimple groups that are not semidirect products, our survey of extensions is still incomplete. Notice the sort of survey we already have; if we know K, Q, and θ, then we know $K \times_\theta Q$ in the sense that we can write a multiplication table for it.

In discussing extensions G of K by Q, it is convenient to use the multiplicative notation for Q and the additive notation for G and its subgroup K (this is one of the few instances in which one uses the additive notation

for a nonabelian group). Thus, if $k \in K$ and $g \in G$, the conjugate of k by g is written $g + k - g$.

Definition If $\pi: G \to Q$ is onto and $x \in Q$, then a **lifting** of x is an element $g \in G$ with $\pi(g) = x$.

Lemma 7.12 *Let G be an extension of K by Q, where K is abelian. There is a homomorphism $\theta: Q \to \mathrm{Aut}(K)$ such that*

$$\theta_x(k) = l(x) + k - l(x)$$

for every $k \in K$ and every lifting $l(x)$ of x.

Proof If $a \in G$, let f_a denote conjugation by a. Since $K \lhd G$, $f_a|K$ is an automorphism of K, and the function $\mu: G \to \mathrm{Aut}(K)$ defined by $\mu(a) = f_a|K$ is easily seen to be a homomorphism. If $a \in K$, then $\mu(a) = 1_K$, for K is abelian. There is thus a homomorphism $\mu_\#: G/K \to \mathrm{Aut}(K)$ defined by $\mu_\#(Ka) = \mu(a)$.

The first isomorphism theorem not only says that Q and G/K are isomorphic; it says that for any choice of liftings $l(x)$, $x \in Q$, the map $\lambda: Q \to G/K$ defined by $\lambda(x) = Kl(x)$ is an isomorphism. Let $\theta: Q \to \mathrm{Aut}(K)$ be the composite $\mu_\#\lambda$. If $x \in Q$ and $l(x)$ is a lifting, then $\theta_x = \mu_\#\lambda(x) = \mu_\#(Kl(x)) = \mu(l(x)) \in \mathrm{Aut}(K)$; therefore, if $k \in K$,

$$\theta_x(k) = \mu(l(x))(k) = l(x) + k - l(x). \quad \blacksquare$$

There is a version of Lemma 7.12 that provides a homomorphism θ when K is not assumed abelian (then $\theta: Q \to \mathrm{Aut}(K)/\mathrm{Inn}(K)$), but this general situation is rather complicated. From now until we discuss the Schur-Zassenhaus lemma, we assume that K is abelian.

We shall examine (not necessarily split) extensions given the data: abelian K, arbitrary Q, and $\theta: Q \to \mathrm{Aut}(K)$. This study is called **cohomology of groups** because of an analogy with methods of cohomology theory in algebraic topology. Our aim is to be able to write a multiplication table (rather, an addition table!) for any extension G arising from the data K, Q, and θ.

Once a homomorphism $\theta: Q \to \mathrm{Aut}(K)$ is given, we abbreviate the expression $\theta_x(k)$ by xk. The following formulas are valid for all $x, y, 1 \in Q$ and $k, k' \in K$:

$$\begin{aligned} x(k + k') &= xk + xk'; \\ (xy)k &= x(yk); \\ 1k &= k. \end{aligned}$$

(These formulas should remind the reader of the definition of a module. Indeed, a (possibly noncommutative) ring ZQ can be constructed, the *integral group ring* of Q, and the abelian group K is a left module admitting operators from this ring. This remark is the basis of further examinations of group extensions; for more details we refer the reader to the book of S. Mac Lane, *Homology*, Chapter IV.)

Since the object of our study is to recapture G from K and Q, let us assume we begin with only K, Q, and θ, but that there is no G in sight.

Definition Let K and Q be groups, K abelian, and $\theta: Q \to \mathrm{Aut}(K)$ a homomorphism. An extension G of K by Q **realizes** θ in case

$$xk = l(x) + k - l(x)$$

for every $k \in K$ and every lifting $l(x)$ of x.

In this language, Lemma 7.12 says that every extension G of an abelian K by Q determines a θ that G realizes. The intuitive meaning of θ is that it describes how K is normal in G.

The extension problem is now posed as follows: Given an abelian K, a group Q, and a homomorphism $\theta: Q \to \mathrm{Aut}(K)$, determine all extensions G of K by Q that realize θ. (For a discussion of the more complicated case when K is not abelian, the reader is referred to the book of Mac Lane.)

Let G be an extension of K by Q and let $\pi: G \to Q$ be a homomorphism of G onto Q having kernel K. Choosing a lifting $l(x)$ of each $x \in Q$ defines a function $l: Q \to G$ (which may not be a homomorphism) satisfying $\pi \circ l = 1_Q$. The range of such a function l is called a **transversal** of K in G (or a **complete set of coset representatives** of K in G), for it consists of exactly one representative from each coset of K.

Suppose K is abelian and G is an extension of K by Q realizing θ. We identify Q with G/K, and for each $x \in Q$, we choose a lifting $l(x) \in G$; for computational ease, choose $l(1) = 0$. Once this choice of transversal has been made, every element $g \in G$ has a unique expression of the form

$$g = k + l(x) \qquad x \in Q, k \in K$$

(after all, $l(x)$ is a representative of a coset of K in G, and G is the disjoint union of these cosets). We have the following formulas:

(i) $l(x) + k = xk + l(x)$ for every $x \in Q$ and $k \in K$ (this is the statement that G realizes θ);

(ii) $l(x) + l(y) = f(x, y) + l(xy)$ for some $f(x, y) \in K$ ($l(x) + l(y)$ and $l(xy)$ are representatives of the same coset of K).

Definition The function $f: Q \times Q \to K$ defined by formula (ii) is called a **factor set**[2] of G.

Notice that a factor set depends on the choice of transversal. If G is a semidirect product, Exercise 7.29 shows that there is a transversal l making $f(x, y) = 0$ for every x and y in Q, namely, $l(x) = (0, x)$, all $x \in Q$. In fact, it follows from property (ii) that $l: Q \to G$ is a homomorphism if and only if the corresponding factor set is identically zero. Thus, a factor set may fruitfully be thought of as a "measure" of G's deviation from being a semidirect product, for it tells how the transversal l fails to be a homomorphism.

Theorem 7.13 *Let K be abelian and let $\theta: Q \to \mathrm{Aut}(K)$. A function $f: Q \times Q \to K$ is a factor set if and only if it satisfies the formulas:*

(i) $f(1, y) = 0 = f(x, 1)$ *for every* $x, y \in Q$;

(ii) $xf(y, z) - f(xy, z) + f(x, yz) - f(x, y) = 0$ *for every* $x, y, z \in Q$.

Proof Suppose f is a factor set. This means that there is an extension G of K by Q realizing θ, a transversal l has been chosen, and f satisfies

$$l(x) + l(y) = f(x, y) + l(xy).$$

Since we have assumed that $l(1) = 0$, $0 + l(y) = f(1, y) + l(1\,y)$, so that $f(1, y) = 0$; a similar calculation shows that $f(x, 1) = 0$, so that condition (i) is established.

Formula (ii) arises from associativity; one need only calculate the consequences of $[l(x) + l(y)] + l(z) = l(x) + [l(y) + l(z)]$.

Suppose, conversely, that we have a function $f: Q \times Q \to K$ which satisfies (i) and (ii). We shall construct an extension G of K by Q realizing θ and we shall choose a transversal l such that f is the factor set determined by this data.

Let G be the set of all pairs $(k, x) \in K \times Q$ with the binary operation

$$(k, x) + (k', y) = (k + xk' + f(x, y), xy).$$

(Note the similarity to the construction of $K \times_\theta Q$.)

The proof that G is a group is quite similar to the proof of Theorem 7.9; formula (ii) is needed to prove associativity; the identity is $(0, 1)$; the inverse of (k, x) is $(-x^{-1}k - x^{-1}f(x, x^{-1}), x^{-1})$.

That G is an extension of K by Q is also easy; identify K with all pairs of the form $(k, 1)$ and define $\pi: G \to Q$ by $\pi(k, x) = x$.

[2] The term *factor set* is a misnomer; factor function would be more suggestive. However, we conform to standard usage. Factor sets are often called **cocycles** to stress the analogy with cohomology theory.

Does G realize θ? We must show that if $x \in Q$ and $k \in K$, then $xk = l(x) + k - l(x)$, where $l(x)$ is any lifting of x. We identify xk, as any element of K, with $(xk, 1)$; also the definition of π yields $l(x) = (k', x)$ for some $k' \in K$. Compute:

$$\begin{aligned}(k', x) + (k, 1) - (k', x) &= (k' + xk, x) + (-x^{-1}k', x^{-1}) \\ &= (k' + xk + x(-x^{-1}k'), 1) \\ &= (xk, 1),\end{aligned}$$

for K is abelian.

Finally, define a transversal l by $l(x) = (0, x)$. A straightforward computation shows that $l(x) + l(y) - l(xy) = (f(x, y), 1)$, as desired. ∎

Definition $Z_\theta^2(Q, K)$ is the set of all factor sets $f: Q \times Q \to K$.

EXERCISE

7.34. Prove that $Z_\theta^2(Q, K)$ is an abelian group under the operation

$$(f + f')(x, y) = f(x, y) + f'(x, y).$$

We are near a solution of the extension problem. Given abelian K, Q, and θ, we can write a multiplication table for an extension G of K by Q realizing θ once we are given a factor set. Moreover, factor sets are concrete objects that are succinctly characterized by Theorem 7.13.

It is quite possible that different factor sets f and f' give rise to the same extension. Indeed, let us again assume that we have an extension G, and recall that a factor set f is defined by

$$l(x) + l(y) = f(x, y) + l(xy).$$

The factor set f thus depends on the transversal l. If we choose a different transversal l', we get a second factor set f', defined by

$$l'(x) + l'(y) = f'(x, y) + l'(xy),$$

but, quite clearly, f' still gives the group G.

Lemma 7.14 *Let G be an extension of K by Q realizing θ; let l and l' be transversals giving rise to factor sets f and f'. There exists a function $\alpha: Q \to K$ with $\alpha(1) = 0$ and*

$$f'(x, y) - f(x, y) = x\alpha(y) - \alpha(xy) + \alpha(x)$$

for every $x, y \in Q$.

Proof For each $x \in Q$, $l(x)$ and $l'(x)$ are just different representatives of the same coset of K in G. There is thus an element $\alpha(x) \in K$ with

$$l'(x) = \alpha(x) + l(x).$$

Since we are consistently lifting 1 to 0, $l'(1) = l(1) = 0$, so that $\alpha(1) = 0$. The main formula is derived as follows:

$$\begin{aligned} l'(x) + l'(y) &= \alpha(x) + l(x) + \alpha(y) + l(y) \\ &= \alpha(x) + x\alpha(y) + l(x) + l(y) \qquad \text{(since } G \text{ realizes } \theta) \\ &= \alpha(x) + x\alpha(y) + f(x, y) + l(xy) \\ &= \alpha(x) + x\alpha(y) + f(x, y) - \alpha(xy) + l'(xy). \end{aligned}$$

It follows that $f'(x, y) = \alpha(x) + x\alpha(y) + f(x, y) - \alpha(xy)$. This gives the desired formula, for each term lies in the abelian group K. ∎

Definition A **coboundary** is a function $g: Q \times Q \to K$ such that

$$g(x, y) = x\alpha(y) - \alpha(xy) + \alpha(x)$$

for some $\alpha: Q \to K$ with $\alpha(1) = 0$. $B^2_\theta(Q, K)$ is the set of all coboundaries.

EXERCISE

7.35. $B^2_\theta(Q, K)$ is a subgroup of $Z^2_\theta(Q, K)$.

Lemma 7.14 says that factor sets arising from different transversals lie in the same coset of B^2 in Z^2. We have been led to the following quotient group and equivalence relation.

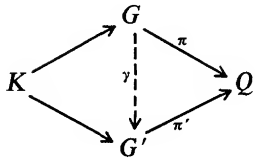
Definition $H^2_\theta(Q, K)$ is defined to be $Z^2_\theta(Q, K)/B^2_\theta(Q, K)$ and is called the **second cohomology group**.

Definition Two extensions G and G' of K by Q realizing θ are **equivalent** if there are factor sets f of G and f' of G' with $f' - f \in B^2_\theta(Q, K)$.

Observe that we have defined a relation on groups, not merely on factor sets: If one chooses other factor sets f_0 of G and f'_0 of G', then $f'_0 - f_0 = (f'_0 - f') - (f_0 - f) + (f' - f) \in B^2_\theta(Q, K)$, since $f'_0 - f'$ and $f_0 - f$ are in $B^2_\theta(Q, K)$, by Lemma 7.14.

EXERCISES

***7.36.** Let G and G' be extensions of K by Q realizing θ, where K is abelian. We have the diagram



where $\pi: G \to Q$ and $\pi': G' \to Q$ are onto. Prove that G is equivalent to G' if and only if there is an isomorphism $\gamma: G \to G'$ such that $\gamma(k) = k$ for all $k \in K$ and $\pi' \circ \gamma = \pi$. (HINT: Use multiplication tables for G and G' to define γ.)

7.37. Let K be abelian and let $\theta: Q \to \text{Aut}(K)$. Prove that any two semidirect products of K by Q realizing θ are equivalent.

7.38. Let p be an odd prime. Give an example of isomorphic extensions of $\sigma(p)$ by $\sigma(p)$ which are not equivalent.

We summarize the results of this section in the following theorem.

Theorem 7.15 *Let K be abelian and let $\theta: Q \to \text{Aut}(K)$. The set E of all equivalence classes of extensions of K by Q realizing θ forms an abelian group isomorphic to $H^2_\theta(Q, K)$ whose identity element is the class of the semidirect product.*

Proof If G is an extension of K by Q realizing θ, let $[G]$ denote its equivalence class, so that E is the set of all $[G]$. We now define a one-to-one correspondence $\lambda: H^2_\theta(Q, K) \to E$ as follows: $\lambda(f + B^2) = [G_f]$, where G_f is the extension determined by f that we constructed in Theorem 7.13. First of all, λ is well defined, for if both f and f' lie in the same coset of B^2, the definition of equivalence says that G_f and $G_{f'}$ are equivalent. Conversely, λ is one-to-one, for if $\lambda(f + B^2) = \lambda(f' + B^2)$, then G_f and $G_{f'}$ are equivalent, and so $f' - f \in B^2$. Furthermore, λ is onto, for if G is an extension and f is a factor set of G, then $\lambda(f + B^2) = [G]$.

By Exercise 1.19, there is a unique addition defined on E making it a group and λ an isomorphism. The last part of the theorem follows from the fact that an extension is a semidirect product if and only if it has a factor set in B^2. ∎

In Chapter 10, we shall give an explicit construction of the sum of two classes in E.

Corollary 7.16 *Let K be abelian and let $\theta: Q \to \text{Aut}(K)$. Every extension of K by Q realizing θ is a semidirect product if and only if $H^2_\theta(Q, K) = \{0\}$.*

THE SCHUR-ZASSENHAUS LEMMA

We now apply the results of the preceding section.

Theorem 7.17 *Let K be an abelian group of order m and let Q be a group of order n. If $(m, n) = 1$, then every extension G of K by Q is a semidirect product.*

Proof By Corollary 7.16, it suffices to prove that every factor f lies in $B_0^2(Q, K)$.

Define $\sigma: Q \to K$ by

$$\sigma(x) = \sum_{y \in Q} f(x, y);$$

σ is well defined, since Q is finite and K is abelian. Now, if we sum the formula

$$xf(y, z) - f(xy, z) + f(x, yz) = f(x, y)$$

over all $z \in Q$, we obtain

$$x\sigma(y) - \sigma(xy) + \sigma(x) = nf(x, y)$$

(as z ranges over all elements of Q, so does yz for fixed y). Since $(m, n) = 1$, there are integers a and b with $am + bn = 1$. If we define $\alpha: Q \to K$ by $\alpha(x) = b\sigma(x)$, then $\alpha(1) = 0$ and

$$x\alpha(y) - \alpha(xy) + \alpha(x) = f(x, y).$$

This formula says that $f \in B_0^2(Q, K)$, as desired. ∎

Theorem 7.18 *Let K be an abelian group of order m, Q a group of order n, $(m, n) = 1$, and G an extension of K by Q. Then any two subgroups of G of order n are conjugate.*

Proof Let Q_1 be a subgroup of G of order n. Now $K \cap Q_1 = \{1\}$, for if $a \in K \cap Q_1$, its order must divide m and n, and $(m, n) = 1$. Moreover, $|KQ_1| = |KQ_1||K \cap Q_1| = |K||Q_1| = |G|$, so that $KQ_1 = G$. It follows that G is a semidirect product of K by Q_1. If Q_2 is a second subgroup of G of order n, then G is a semidirect product of K by Q_2 as well. By Exercise 7.29, Q_1 consists of liftings $l_1(x)$, and Q_2 consists of liftings $l_2(x)$, where l_1 and l_2 are homomorphisms; it follows that the factor sets f_1 and f_2 determined by the $l_1(x)$ and by the

$l_2(x)$ are each zero. Now, if $l_1(x) = \alpha(x) + l_2(x)$, then

$$0 = f_1(x, y) - f_2(x, y) = x\alpha(y) - \alpha(xy) + \alpha(x).^3$$

Summing over all $y \in Q$ gives the equation:

$$0 = xk - k + n\alpha(x)$$

where

$$k = \sum_{y \in Q} \alpha(y).$$

This is an equation in K. If $am + bn = 1$ and $h = bk$,

$$-\alpha(x) = xh - h = -h + xh.$$

It follows that $-h + Q_1 + h = Q_2$, for $-h + l_1(x) + h = -h + xh + l_1(x) = -\alpha(x) + l_1(x) = l_2(x) - l_1(x) + l_1(x) = l_2(x)$. ∎

Theorem 7.19 (Schur-Zassenhaus Lemma) *If K and Q are finite groups of orders m and n, respectively, and if $(m, n) = 1$, then every extension G of K by Q is a semidirect product.*

Proof As we saw at the beginning of the proof of Theorem 7.18, it suffices to prove that G contains a subgroup of order n.

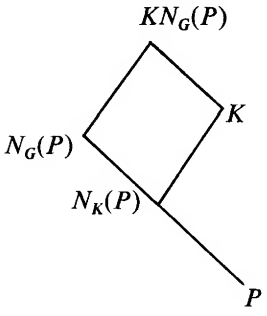
We perform an induction on m, noting that if $m = 1$, the theorem is trivial. Suppose K contains a proper subgroup T that is also normal in G. Then $K/T \lhd G/T$ and $(G/T)/(K/T) \cong G/K \cong Q$. If $|K/T| = m'$, then $m' < m$ and $|G/T| = m'n$. By induction, G/T contains a subgroup N/T of order n. Now $|N| = n|T|$ and $(n, |T|) = 1$, for $|T|$ divides $|K| = m$. Since $T \lhd N$ and $|T| < m$, the inductive hypothesis gives a subgroup of N, and hence of G, of order n.

We may now assume that K is a minimal normal subgroup of G. If p is a prime dividing m and if P is a p-sylow subgroup of K, then a check of its order shows that P is also a p-sylow subgroup of G, since $(m, n) = 1$. Furthermore, since $K \lhd G$, P has the same number of conjugates in G as in K, so that $[G: N_G(P)] = [K: N_K(P)]$. By Lagrange's theorem,

$$[N_G(P): N_K(P)] = [G: K] = n.$$

[3] A function $\alpha: Q \to K$ such that $\alpha(xy) = \alpha(x) + x\alpha(y)$ (as this one) is called a **crossed homomorphism.**

Since $N_K(P) = K \cap N_G(P)$, the second isomorphism theorem gives $[KN_G(P): K] = [N_G(P): N_K(P)] = n$, so that $|KN_G(P)| = n|K| = |G|$.



Since $KN_G(P)$ is a subgroup of G, it follows that $KN_G(P) = G$. Applying the second isomorphism theorem once more gives

$$N_K(P) = K \cap N_G(P) \lhd N_G(P)$$

and

$$\begin{aligned} N_G(P)/N_K(P) &= N_G(P)/(K \cap N_G(P)) \\ &\cong KN_G(P)/K = G/K \cong Q. \end{aligned}$$

If $N_G(P)$ is a proper subgroup of G, then $|N_K(P)| < m$, and so $N_G(P)$ contains a subgroup of order n, by induction. Therefore, we may assume that $N_G(P) = G$, i.e., that $P \lhd G$.

Since $K \supset P$ and K is a minimal normal subgroup of G, $K = P$. Now $Z(P)$ is a characteristic subgroup of P, so that $P \lhd G$ implies $Z(P) \lhd G$. Again, minimality implies $Z(P) = P$, for a finite p-group has a nontrivial center. But now P is abelian, and the proof is completed by Theorem 7.17. ∎

The generalization of Theorem 7.18 is true: Under the hypotheses of the Schur-Zassenhaus lemma, any two subgroups of order n are conjugate. One proves this theorem by first assuming that either K or Q is solvable. Since $(m, n) = 1$, one of these groups must have odd order, and hence one of K or Q must be solvable, by the theorem of Feit and Thompson (see Exercise 6.27).

Had we known Theorems 7.17 and 7.18, we could have applied them in the proof of P. Hall's theorem (Theorem 6.19), for the second case in that proof is the situation in which the normal subgroup is abelian.

EXERCISES

7.39. Use the Schur-Zassenhaus lemma and Exercise 7.26 to reclassify all groups of order pq, where p and q are distinct primes.

7.40. Prove that every group of order p^2q, where $p > q$ are primes, has a normal p-sylow subgroup, and classify all such groups.

7.41. Using Corollary 5.17, reclassify all groups of order 12.

7.42. Use factor sets to prove the existence of the generalized quaternions Q_n. (HINT: Use Exercise 5.50.)

7.43. Let $p_1 < p_2 \cdots < p_t$ be primes such that $(p_i, p_j - 1) = 1$ for all $i < j$. Prove that every group G of order $n = p_1 p_2 \cdots p_t$ is cyclic. (HINT: Use Exercise 5.55, which states that G must contain a normal p_t-sylow subgroup.)

We sketch a proof that there exist sets of primes $\{p_1, p_2, \cdots, p_t\}$ for arbitrarily large t that satisfy the hypothesis of the last exercise. Let $p_1 = 3$, and suppose there are $p_1 < p_2 < \cdots < p_t$ with $(p_i, p_j - 1) = 1$ for all $i < j$. We quote a theorem of Dirichlet: If $(a, b) = 1$, then the arithmetic progression $a, a + b, a + 2b, \cdots, a + mb, \cdots$ contains infinitely many primes. Since $(2, p_1 p_2 \cdots p_t) = 1$, there is a positive integer m such that $p_{t+1} = 2 + mp_1 p_2 \cdots p_t$ is prime; $\{p_1, p_2, \cdots, p_{t+1}\}$ satisfies the desired conditions.

The reader may test his group-theoretical muscles on the following problem: For which integers n is every group of order n abelian? The number $5929 = 7^2 11^2$ should be considered before any false conjectures are made.

CHAPTER 8

Some Simple Groups

FINITE FIELDS

The Jordan-Hölder theorem tells us that once we know extensions and simple groups, we know all groups that possess composition series. Several families of simple groups are known (in addition to the cyclic groups of prime order and the large alternating groups), and our main concern in this chapter is the exhibition of one of these families, the projective unimodular groups. These are essentially groups of matrices of determinant 1 whose entries are allowed to lie in (almost) any field. Since these groups are finite only when the underlying field is finite, we begin our discussion by examining the finite fields.

Definition A **prime field** is a field having no proper subfields.

Theorem 8.1 *Every prime field K is isomorphic to either Z_p or the rationals,* $\mathbf{Q}$.

Proof If 1 is the unit element of K, let $R = \{n \cdot 1 : n \in Z\} \subset K$; it is easily checked that R is a domain. Further, the map $f: Z \to R$ defined by $f(n) = n \cdot 1$ is a ring homomorphism of Z onto R. By the first isomorphism theorem for rings, $R \cong Z/I$, where I is the kernel of f. Since R is a domain, I is a prime ideal. If $I = (p)$, then $R \cong$

$Z/(p) \cong Z_p$. In this case, R is a field, and so $R = K$. If $I = (0)$, then $R \cong Z$; it follows that K contains a subfield S isomorphic to the rationals, for K must contain the multiplicative inverse of each nonzero element in R. Since K is a prime field, $K = S$. ∎

Theorem 8.2 *Every field F contains a unique prime field K.*

Proof Let K be the intersection of all subfields of F; recall that every subfield of F contains 1, so that $K \neq \{0\}$. The reader may now prove that K is a prime field that is visibly the only prime field contained in F. ∎

Definition Let F be a field with prime field K; F has **characteristic p** if $K \cong Z_p$; otherwise, F has **characteristic 0** (in which case $K \cong \mathbf{Q}$).

Observe that if F has characteristic p, then $pa = 0$ for every $a \in F$.

Corollary 8.3 *If F is a finite field, then F has exactly p^n elements for some prime p.*

Proof Since F is finite, it must have characteristic p for some prime p. Moreover, F is a finite-dimensional vector space over Z_p. If F has dimension n, then one may choose a basis of F and count exactly p^n vectors $(k_1, k_2, \cdots, k_n)$, where $k_i \in Z_p$. ∎

We remark that there do exist infinite fields of characteristic p, e.g., all rational functions over Z_p (a rational function is a quotient of two polynomials).

The next question is whether there are any finite fields aside from the Z_p. In order to exhibit them, we must generalize the method of adjoining a root of a polynomial to a field. In Chapter 6, this process was simple because we could work within the complex numbers. We must be wily here, however, for there is no such larger field available to us at the outset.

Lemma 8.4 *Let K be a field and let $\varphi(x)$ be an irreducible polynomial in $K[x]$. There exists a field F containing a root of $\varphi(x)$ and a subfield isomorphic to K.*

Proof Let $R = K[x]$ and let I be the ideal $(\varphi(x))$. Since $\varphi(x)$ is irreducible, R/I is a field, by Corollary 4.20.

Set $F = R/I$. If $\varphi(x) = \sum k_i x^i$, where $k_i \in K$, then

$$\begin{aligned}\varphi(x + I) &= \sum k_i(x + I)^i \\ &= \sum k_i(x^i + I) \\ &= \sum (k_i x^i + I) \\ &= (\sum k_i x^i) + I \\ &= \varphi(x) + I.\end{aligned}$$

Since $\varphi(x) \in I$, $\varphi(x) + I = I$, the zero element of F. Therefore, the element $x + I$ in F is a root of $\varphi(x)$.

Finally, the reader may verify that $\{k + I: k \in K\}$ is a subfield of F that is isomorphic to K. ∎

Theorem 8.5 *If K is a field and $f(x) \in K[x]$, there is a field F containing K over which $f(x)$ is a product of linear factors.*

Proof We perform an induction on the degree d of $f(x)$. If $d = 1$, then K itself is the desired field. If $d > 1$, then $f(x) = \varphi(x)g(x)$, where $\varphi(x)$ is an irreducible polynomial. By Lemma 8.4, there is a field F_0 containing K and a root α of $\varphi(x)$. In $F_0[x]$, $\varphi(x) = (x - \alpha)\psi(x)$, and so $f(x) = (x - \alpha)\psi(x)g(x)$. Since $\psi(x)g(x)$ has degree less than d, the inductive hypothesis provides a field F containing F_0 (hence containing K and α) in which $\psi(x)g(x)$ (hence $f(x)$) is a product of linear factors. ∎

EXERCISES

**8.1. If F is a field of characteristic $p > 0$, then for every $a, b \in G$,

$$(a + b)^{p^k} = a^{p^k} + b^{p^k} \qquad \text{for all } k > 0.$$

**8.2. If F is a finite field with exactly q elements, then every element in F is a root of $x^q - x$. (HINT: Use Lagrange's theorem on the nonzero elements of F.)

Definition Let F be a field and let

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 \in F[x].$$

The **derivative** of $f(x)$, denoted $f'(x)$, is the polynomial

$$f'(x) = na_nx^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1 \in F[x].$$

EXERCISES

8.3. Prove the usual formulas of calculus for the derivatives of sums and products of polynomials in $F[x]$.

****8.4.** Let $f(x) \in F[x]$. Then $f(x)$ and $f'(x)$ have a nonconstant common factor if and only if $f(x)$ has repeated roots.

Theorem 8.6 *Let p be a prime and n be a positive integer. There exists a field F with exactly p^n elements.*

Proof Let $q = p^n$, and consider the polynomial

$$f(x) = x^q - x \in Z_p[x].$$

By Theorem 8.5, there is a field F_0 containing Z_p in which $f(x)$ is a product of linear factors. Note that since F_0 contains Z_p, it has characteristic p.

Let F be the subset of F_0 consisting of all the roots of $f(x)$. Since $f(x)$ has degree q, $|F| \leq q$. To prove that $|F| = q$, it suffices to prove that all the roots of $f(x)$ in F_0 are distinct, i.e., that $f(x)$ has no repeated roots. Now $f'(x) = qx^{q-1} - 1 = -1$, since $q = p^n$ and F_0 has characteristic p. Now apply Exercise 8.4.

Let us prove F is a field. If a and b are roots of $f(x)$, then, using Exercise 8.1, $(a-b)^q = a^q - b^q = a - b$, i.e., $a - b$ is a root of $f(x)$. Hence, F is a group under addition. Further, $(ab)^q = a^qb^q = ab$, so that ab is also a root of $f(x)$; thus, F is commutative ring. Finally, if $f(a) = 0$ and $a \neq 0$, then $a^{q-1} = 1$; the inverse of a is thus a^{q-2}. Therefore, F is a field having precisely $q = p^n$ elements. ∎

Theorem 8.7 *Any two fields having exactly p^n elements are isomorphic.*

Proof If $|F| = p^n = q$, then, by Exercise 8.2, every element in F is a root of $f(x) = x^q - x$. In the language of Chapter 6, F is a root field of $f(x)$. The reader may now adapt the proof of Lemma 6.4 to prove its analog: Any two root fields of $f(x)$ over Z_p are isomorphic. The result follows. ∎

It was Galois who discovered these finite fields, and they are called **Galois fields** in his honor. The field with p^n elements is thus denoted $GF(p^n)$.

THE GENERAL LINEAR GROUP

Groups of nonsingular matrices are as natural an object of study as groups of permutations. In investigating the structure of these groups, we shall discover a new family of simple groups.

Definition Let K be a field. The **general linear group** $GL(m, K)$ is the multiplicative group of all nonsingular $m \times m$ matrices over K.

If $K = GF(q)$, one usually denotes this group $GL(m, q)$. A slight generalization of the result of Exercise 7.8 is that $GL(m, q) \cong \text{Aut}(V)$, where V is a direct sum of m copies of $\sigma(q)$.

Theorem 8.8 $|GL(m, q)| = (q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})$.

Proof Let V be an m-dimensional vector space over $GF(q)$, and let $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ be an ordered basis. If we regard $GL(m, q)$ as linear transformations on V, then we can exhibit a one-to-one correspondence between $GL(m, q)$ and the family of all ordered bases of V. If $T \in GL(m, q)$, then $\{T\alpha_1, T\alpha_2, \cdots, T\alpha_m\}$ is an ordered basis of V (because T is nonsingular); if $\{\beta_1, \beta_2, \cdots, \beta_m\}$ is an ordered basis of V, there is a unique $T \in GL(m, q)$ with $T\alpha_i = \beta_i$ for all i.

An ordered basis of V consists of vectors $\{\beta_1, \beta_2, \cdots, \beta_m\}$. Since there are q^m vectors in V, there are $q^m - 1$ choices for β_1 (the zero vector is not a candidate). Having chosen β_1, the only restriction on β_2 is that it not lie in the subspace spanned by β_1; there are thus $q^m - q$ choices for β_2. More generally, having chosen an independent set $\{\beta_1, \beta_2, \cdots, \beta_i\}$, the only restriction on β_{i+1} is that it not lie in the subspace spanned by $\{\beta_1, \beta_2, \cdots, \beta_i\}$; there are thus $q^m - q^i$ choices for β_{i+1}. Therefore, there are exactly $(q^m - 1) \times (q^m - q) \cdots (q^m - q^{m-1})$ ordered bases of V. ∎

Notation If K is a field, then K^* is the multiplicative group of its nonzero elements.

Definition A **primitive element** of $GF(q)$ is a generator of $GF(q)^*$ (we saw in Exercise 4.17 that $GF(q)^*$ is a cyclic group).

Let ρ be a primitive element of $GF(q)$.

Definition If t is a nonnegative integer, then

$$M(t) = \{A \in GL(m, q): \det A = \text{power of } \rho^t\}.$$

Lemma 8.9 *Let t be a divisor of $q - 1$. If $\Omega = |GL(m, q)|$, then $M(t)$ is a normal subgroup of $GL(m, q)$ of order Ω/t.*

Proof We use the correspondence theorem in the setting

$$\det: GL(m, q) \to GF(q)^*.$$

If t is a divisor of $q - 1 = |GF(q)^*|$, then the cyclic subgroup $\langle \rho^t \rangle$ of $GF(q)^*$ has order $(q - 1)/t$, and hence index t. Furthermore, $\langle \rho^t \rangle \lhd GF(q)^*$, for the latter group is abelian. Since $M(t)$ is the subgroup of $GL(m, q)$ that corresponds to $\langle \rho^t \rangle$, we have $M(t) \lhd GL(m, q)$, of index t, and of order Ω/t. ∎

Theorem 8.10 *Let $q - 1 = p_1 p_2 \cdots p_k$, where the p are (not necessarily distinct) primes. The following normal series is the beginning of a composition series:*

$$GL(m, q) = M(1) \supset M(p_1) \supset M(p_1 p_2) \supset \cdots \supset M(q - 1).$$

Proof We have already seen above that each of the terms in this series is normal in $GL(m, q)$. Furthermore, if $\Omega = |GL(m, q)|$,

$$\left| \frac{M(p_1 p_2 \cdots p_i)}{M(p_1 p_2 \cdots p_{i+1})} \right| = \frac{\Omega/p_1 p_2 \cdots p_i}{\Omega/p_1 p_2 \cdots p_{i+1}} = p_{i+1}.$$

Since the factor groups have prime order, they are simple. ∎

The last subgroup in the above chain, $M(q - 1)$, is of special interest; it consists of all matrices of determinant $\rho^{q-1} = 1$.

Definition A matrix is **unimodular** if it has determinant 1.

Definition Let K be a field. The **special linear group** $SL(m, K)$ is the multiplicative group of all $m \times m$ unimodular matrices over K.

If $K = GF(q)$, one usually denotes this group $SL(m, q)$; it has order $\Omega/q - 1$.

One would discover the subgroup SL without recalling determinants, for SL is the commutator subgroup of GL, as we shall see later.

The following elementary matrices are introduced in order to analyze the structure of the subgroup $SL(m, K)$.

Definition Let λ be a nonzero element of K and $i \neq j$ integers between 1 and m. A **transvection**[1] $B_{ij}(\lambda)$ is the $m \times m$ matrix differing from the identity matrix E in that it has λ as its ijth entry.

Note that every transvection is unimodular. Also, $B_{ij}(\lambda)A$ is the matrix obtained from A by adding λ times its jth row to its ith row.

Lemma 8.11 *If $A \in GL(m, K)$, then $A = UD(\mu)$, where U is a product of transvections and $D(\mu)$ is the diagonal matrix with diagonal entries $\{1, 1, \cdots, 1, \mu\}$.*

Proof We prove by induction on t that A can be transformed, by a sequence of operations that add a multiple of one row to another row, into a matrix of the form

$$A' = \begin{bmatrix} E_t & * \\ 0 & C \end{bmatrix},$$

where E_t is a $t \times t$ identity matrix.

Since A is nonsingular, not every entry a_{i1} in the first column is 0. Adding some row to the second if necessary, we may assume $a_{21} \neq 0$. Now add $a_{21}^{-1}(1 - a_{11})$ times the second row to the first row to get 1 in the 1,1 position. We may now clean out the remainder of the first column so that all other entries in it are 0.

Suppose A has been transformed into A'. Note that A' is nonsingular, so that the matrix C is also nonsingular. As before, we may insert 1 in the $t + 1, t + 1$ position, using row operations involving only the rows of C (assuming that C has at least two rows); it follows that the first t columns of A' are not altered by this insertion. Adding on a suitable multiple of row $t + 1$ to the other rows cleans out column $t + 1$ so that all its entries are 0 except the 1 in position $t + 1, t + 1$.

This process may be continued until A has the form

$$\begin{bmatrix} E_{n-1} & * \\ 0 & \mu \end{bmatrix},$$

where $\mu \neq 0$. Adding suitable multiples of the last row to the other rows cleans out the last column, leaving $D(\mu)$.

[1] Most authors call every matrix similar to some $B_{ij}(\lambda)$ a transvection.

In terms of matrix multiplication, we have shown that $PA = D(\mu)$, where P is a product of transvections. Since the inverse of $B_{ij}(\lambda)$ is $B_{ij}(-\lambda)$, which is again a transvection, $A = P^{-1}D(\mu)$ is the desired factorization. ∎

Theorem 8.12 *$GL(m, K)$ is a semidirect product of $SL(m, K)$ by K^*; $SL(m, K)$ is generated by transvections.*

Proof First of all, $SL(m, K) \lhd GL(m, K)$. If Δ is the set of all diagonal matrices $D(\mu)$ (notation as in the lemma), then Δ is a subgroup of $GL(m, K)$ isomorphic to K^*. Since $\det(D(\mu)) = \mu$, $\Delta \cap SL(m, K) = \{E\}$, so that GL is a semidirect product of SL by K^*.

If $A = UD(\mu)$ is a factorization of an element A of GL, as in the lemma, then $\det A = \mu$. It follows that if A is unimodular, $D(\mu) = D(1) = E$, and so $A = U$ is a product of transvections. ∎

Theorem 8.13 *The center of $SL(m, K)$, which we denote Z_0, consists of all scalar matrices kE with $k^m = 1$.*

Proof It suffices to prove that $A \in Z_0$ implies A is a scalar matrix kE; that $k^m = 1$ will then follow from $\det A = 1$.

Let us consider $SL(m, K)$ as transformations of a vector space over K with basis $\{e_1, \cdots, e_m\}$. The transvection $B_{ij}(1)$ is that transformation B which sends e_j into $e_i + e_j$ and which fixes the other basis elements. Let A be the transformation sending e_l into $\sum a_{kl}e_k$.

If $l \neq j$, then $ABe_l = Ae_l$. On the other hand,

$$\begin{aligned} BAe_l = B(\textstyle\sum a_{kl}e_k) &= \textstyle\sum a_{kl}Be_k \\ &= \Big(\sum_{k \neq j} a_{kl}e_k\Big) + a_{jl}(e_i + e_j) \\ &= Ae_l + a_{jl}e_i. \end{aligned}$$

If $AB = BA$, then $a_{jl} = 0$ whenever $j \neq l$; therefore A is diagonal: $Ae_l = a_{ll}e_l$. An easy computation shows that $ABe_j = a_{ii}e_i + a_{jj}e_j$ while $BAe_j = a_{jj}(e_i + e_j)$. Hence $a_{ii} = a_{jj}$. Therefore, if A commutes with all $B_{ij}(1)$, then A is scalar. ∎

Theorem 8.14 *If Z_0 is the center of $SL(m, q)$, then $|Z_0| = d$, where $d = (m, q - 1)$.*

Proof Suppose that $\langle \rho \rangle$ is a cyclic group of order n and that $k \in \langle \rho \rangle$ satisfies $k^m = 1$. If $d = (n, m)$, then $m = dm'$, where $(n, m') = 1$.

There are integers a and b with $na + m'b = 1$, so that

$$k = k^{na+m'b} = k^{na}k^{m'b} = k^{m'b}.$$

Therefore $k^d = k^{m'bd} = k^{mb} = 1$. We have shown that $k^m = 1$ if and only if $k^d = 1$. It now follows easily that $k^m = 1$ if and only if $k \in \langle \rho^{n/d} \rangle$, a group of order d. This proves the theorem, for let $n = q - 1$ and ρ be a primitive element of $GF(q)$. ▮

EXERCISE

**8.5. Let $H \lhd SL(2, K)$ and let $A \in H$. If A is similar to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then there is a nonzero $\mu \in K$ such that H contains $\begin{bmatrix} a & \mu^{-1}b \\ \mu c & d \end{bmatrix}$.

(HINT: Use the factorization of Lemma 8.11.)

PSL(2, K)

Our preceding discussion allows us to extend the normal series of Theorem 8.10 one step further:

$$GL(m, K) \supset \cdots \supset SL(m, K) \supset Z_0.$$

Moreover, Z_0 is an abelian group, so its composition factors are no secret if it is finite. Let us investigate the last factor group of this series.

Definition The **projective unimodular group** $PSL(m, K)$ is the group $SL(m, K)/Z_0$.

If $K = GF(q)$, one usually writes $PSL(m, q)$.

Theorem 8.15 *If* $d = (m, q - 1)$, *then*

$$|PSL(m, q)| = \frac{(q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})}{d(q - 1)}.$$

Proof If $\Omega = |GL(m, q)|$, then $|SL(m, q)| = \Omega/q - 1$, by Lemma 8.9. The theorem now follows from Theorems 8.8 and 8.14. ▮

We first show that $PSL(m, K)$ is of intrinsic interest.

EXERCISE

8.6. If K is a field, let $PLF(K)$ denote the set of all unimodular linear fractional transformations

$$f(x) = \frac{ax + b}{cx + d},$$

where $a, b, c, d \in K$ and $ad - bc = 1$. Prove that $PLF(K)$ is a group (the binary operation is composition of functions) isomorphic to $PSL(2, K)$. (This construction can be generalized to the general case $m \geq 2$.)

If $V_{n+1}(K)$ is an $(n + 1)$-dimensional vector space over a field K, define an equivalence relation on $V_{n+1}(K)$ by $\alpha \equiv \beta$ in case there is a nonzero $k \in K$ with $k\alpha = \beta$; denote the equivalence class of α by $[\alpha]$. Define n-dimensional **projective space** $P_n(K)$ to be the set of all such equivalence classes $[\alpha]$, where $\alpha \neq 0$. Every nonsingular linear transformation $T: V_{n+1}(K) \to V_{n+1}(K)$ defines a function $T^*: P_n(K) \to P_n(K)$ by

$$T^*[\alpha] = [T\alpha];$$

T^* is called a **projective transformation**. If T is unimodular, then T^* is called a **projective unimodular transformation**.

EXERCISE

8.7. The set of all projective unimodular transformations of $P_n(K)$ forms a group isomorphic to $PSL(n+1, K)$.

For the remainder of this section, we concentrate on the case $m = 2$ with the aim of proving that the groups $PSL(2, q)$ are simple when $q > 3$.

Lemma 8.16 *If a normal subgroup H of $SL(2, q)$ contains a transvection $B_{ij}(\mu)$, then $H = SL(2, q)$.*

Proof By Theorem 8.12, it suffices to prove that H contains every transvection.[2] Denote $GF(q)$ by K.

If we conjugate $B_{12}(\mu)$ by a unimodular matrix, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & \mu \\ 0 & 1 \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1 - \mu ac & \mu a^2 \\ -\mu c^2 & 1 + \mu ac \end{bmatrix}.$$

[2] Any two transvections are similar, hence are conjugate in GL; they need not be conjugate in SL.

In particular, this conjugate is $B_{12}(\mu a^2)$ if $c = 0$, and it is $B_{21}(-\mu c^2)$ if $a = 0$. Furthermore, these matrices lie in H, since H is normal.

The map $k \to k^2$ is an endomorphism of the abelian group K^* whose kernel consists of all k with $k^2 = 1$. Since K is a field, the polynomial $x^2 - 1$ has at most two roots, and so the kernel has order 1 or order 2 (it has order 1 if K has characteristic 2). It follows that at least half the elements of K^* are squares.

Let

$$\Gamma = \{\lambda \in K: B_{12}(\lambda) \in H\} \cup \{0\}.$$

It is easy to see that Γ is a subgroup of K (where we consider K only as an additive group). Moreover, we know that Γ contains 0 and all elements of the form μa^2. Therefore, Γ contains more than half the elements of K, and so $\Gamma = K$, by Lagrange's theorem. Hence, H contains all transvections of the form $B_{12}(\lambda)$, and a similar argument shows that H contains all transvections of the form $B_{21}(\lambda)$. ∎

The transvections thus play the same role in the study of the special linear groups as the 3-cycles play in the study of the alternating groups. We now prove the main theorem of this section.

Theorem 8.17 *The groups $PSL(2, q)$ are simple if and only if $q > 3$.*

Proof First of all, Theorem 8.15 gives

$$|PSL(2, q)| = \begin{cases} (q + 1)(q^2 - q) & \text{if } q = 2^n; \\ \frac{1}{2}(q + 1)(q^2 - q) & \text{if } q = p^n, \ p \text{ an odd prime.} \end{cases}$$

Therefore, $|PSL(2, 2)| = 6$ and $|PSL(2, 3)| = 12$, so that these groups cannot be simple.

Let H be a normal subgroup of $SL(2, q)$ that contains a matrix not in Z_0. By the correspondence theorem, it suffices to prove that $H = SL(2, q)$.

Suppose H contains a matrix

$$A = \begin{bmatrix} r & 0 \\ s & t \end{bmatrix},$$

where $r \neq \pm 1$. If

$$S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

then H also contains

$$SAS^{-1}A^{-1} = \begin{bmatrix} 1 & 0 \\ 1 - t^2 & 1 \end{bmatrix}.$$

Since $\det A = 1 = rt$, $t \neq \pm 1$ and $1 - t^2 \neq 0$. This last matrix is thus a transvection, and so $H = SL(2, q)$, by Lemma 8.16.

To complete the proof, we have only to produce a matrix in H whose top row is $[r, 0]$ where $r \neq \pm 1$. Let M be a matrix in H that is not in Z_0. Now M is similar to either a diagonal matrix or a matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & x \end{bmatrix},$$

for the only rational canonical forms for 2×2 matrices are a direct sum of two 1×1 companion matrices, i.e., a diagonal matrix, or a 2×2 companion matrix (which has the above form because M is unimodular). In the first event, we have our desired matrix (since $M \notin Z_0$, and M is unimodular). In the second event, H contains, by Exercise 8.5, a matrix

$$C = \begin{bmatrix} 0 & -\mu^{-1} \\ \mu & x \end{bmatrix}.$$

If

$$T = \begin{bmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{bmatrix},$$

H contains the matrix

$$U = TCT^{-1}C^{-1} = \begin{bmatrix} \alpha^{-2} & 0 \\ \mu x(\alpha^2 - 1) & \alpha^2 \end{bmatrix}.$$

We are done if $\alpha^{-2} \neq \pm 1$, i.e., if $\alpha^4 \neq 1$. If $q > 5$, such a nonzero α does exist in $GF(q)$, for the polynomial $z^4 - 1$ has at most four roots in a field. If $q = 4$, then every $\alpha \in GF(4)$ satisfies the equation $\alpha^4 = \alpha$, so if $\alpha \neq 1$, then $\alpha^4 \neq 1$.

Only the case $q = 5$ remains. There are two possibilities. Assume the entry x in C is nonzero. Choose $\alpha \in GF(5)$ ($\cong Z_5$) so that $\alpha^2 - 1 \neq 0$; note that $\alpha^2 = -1$. The lower left corner of U, $\lambda = \mu x(\alpha^2 - 1)$, is nonzero, and so

$$U = \begin{bmatrix} -1 & 0 \\ \lambda & -1 \end{bmatrix}.$$

Hence, $U^2 = B_{21}(-2\lambda) \in H$; therefore H contains a transvection, and we are done. If the entry x is zero, then

$$C = \begin{bmatrix} 0 & -\mu^{-1} \\ \mu & 0 \end{bmatrix} \in H.$$

Therefore H contains

$$\begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -\mu^{-1} \\ \mu & 0 \end{bmatrix} \begin{bmatrix} 1 & -z \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mu z & -\mu z^2 - \mu^{-1} \\ * & * \end{bmatrix}.$$

If we choose $z = 2\mu^{-1}$, then the top row of this last matrix is $[2, 0]$, as desired. ∎

We have exhibited an infinite family of simple groups. Are any of its members distinct from simple groups we already know? Using the formula of Theorem 8.15, we see that both $PSL(2, 4)$ and $PSL(2, 5)$ have order 60. By Exercise 5.46, any two simple groups of order 60 are isomorphic; therefore

$$PSL(2, 4) \cong A_5 \cong PSL(2, 5).$$

If $q = 7$, however, we do get a new simple group, for $|PSL(2, 7)| = 168$, which is neither prime nor $\frac{1}{2}n!$. If we take $q = 8$, we see that there is a simple group of order 504; if we take $q = 11$, we have a simple group of order 660.

EXERCISES

8.8. The order of $PSL(2, 9)$ is 360. Prove that $PSL(2, 9) \cong A_6$.

8.9. If K is a field of characteristic $\neq 2$, then

$$x = \left(\frac{x+1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^2 \qquad \text{for every } x \in K.$$

Use this remark to prove that $PSL(2, K)$ is simple for every, possibly infinite, field K of characteristic not 2. (HINT: The finiteness of K was used only in the consideration of the subgroup Γ in the proof of Lemma 8.16.)

It is known that $PSL(2, K)$ is simple for every infinite field K.

8.10. Prove that $PSL(2, 2) \cong S_3$.

8.11. Is $SL(2, 3) \cong S_4$? Prove that $PSL(2, 3) \cong A_4$. (HINT: Use Theorem 5.16.)

8.12. What is the 2-sylow subgroup of $SL(2, 3)$?

8.13. What are the composition factors of $GL(2, 7)$?

8.14. Prove that the commutator subgroup of $GL(2, q)$ is $SL(2, q)$ when $q > 3$. What are the commutator subgroups of $GL(2, 3)$ and $GL(2, 2)$?

PSL(m, K)

We shall prove in this section that $PSL(m, K)$ is simple for every field K and all $m \geq 3$. As a consequence of this, we shall be able to exhibit two nonisomorphic simple groups having the same finite order. The proof we present is due to E. Artin and is much more elegant than matrix manipulation. We regard $GL(m, K)$ as all nonsingular transformations on an m-dimensional vector space V over K.

Definition A **hyperplane** H in V is a subspace of dimension $m - 1$ (where $\dim V = m$).

If H is a hyperplane in V and $\alpha \in V$, $\alpha \notin H$, then every vector $\beta \in V$ has a unique expression of the form

$$\beta = \mu\alpha + \gamma, \qquad \mu \in K, \gamma \in H$$

(under the natural map $V \to V/H \cong K$, $\alpha \to \bar{\alpha} \neq 0$, so that the scalar multiples of α form a transversal of H in V).

Lemma 8.18 *Let H be a hyperplane in V and let $A \in GL(m, K)$ fix H (pointwise). If $\alpha \in V$, $\alpha \notin H$, then*

$$A(\alpha) = \mu\alpha + \gamma, \qquad \mu \in K, \gamma \in H.$$

Moreover, given any $\beta \in V$,

$$A(\beta) = \mu\beta + \gamma', \qquad \gamma' \in H.$$

Proof The remarks above show that, as any vector in V, $A(\alpha)$ has an expression as displayed. If $\beta \in V$, then

$$\beta = b\alpha + \gamma'', \qquad b \in K, \gamma'' \in H.$$

Hence,

$$\begin{aligned} A(\beta) &= bA(\alpha) + \gamma'' && \text{(for } A \text{ fixes } H\text{)} \\ &= b(\mu\alpha + \gamma) + \gamma'' \\ (*) \qquad &= \mu(b\alpha + \gamma'') + [(1 - \mu)\gamma'' + b\gamma] \\ &= \mu\beta + \gamma' && \text{where } \gamma' \in H. \end{aligned}$$

∎

The scalar $\mu = \mu(A)$ in Lemma 8.18 is thus uniquely determined by any A that fixes a hyperplane pointwise.

Definition Let $A \in GL(m, K)$ fix a hyperplane H pointwise and let $\mu = \mu(A)$. If $\mu \neq 1$, then A is called a **dilatation**; if $\mu = 1$ and $A \neq E$, then A is called a **transvection**.

The next theorem shows that if a transformation A is a transvection, then there is a matrix representing A that is a transvection $B_{ij}(\lambda)$.

Theorem 8.19 *Let $A \in GL(m, K)$ fix a hyperplane H. If A is a dilatation, then A may be represented (relative to a suitable basis of V) by a matrix $D(\mu) = \text{diagonal}\ \{1, \cdots, 1, \mu\}$; if A is a transvection, then A may be represented by $B_{21}(1)$. Moreover, a transvection has no characteristic vectors outside of H.*

Proof Let us first establish notation. Choose $\alpha \in V$, $\alpha \notin H$, and let

$$A(\alpha) = \mu\alpha + \gamma, \qquad \mu \in K, \gamma \in H.$$

Of course, every nonzero vector in H is a characteristic vector of A; are there any others? If $\beta \in V$, then

$$\beta = b\alpha + \gamma'', \qquad b \in K, \gamma'' \in H.$$

Equation (*) shows that

(**) $\qquad A(\beta) = c\beta$ if and only if $\mu = c$ and $b\gamma = (\mu - 1)\gamma''$.

If A is a dilatation, then $\mu - 1 \neq 0$ and we may solve for

$$\gamma'' = b(\mu - 1)^{-1}\gamma.$$

Therefore, any choice of $b \neq 0$ gives a characteristic vector β outside H. If $\{\tau_1, \cdots, \tau_{m-1}\}$ is a basis of H and if β is a characteristic vector outside H, then $\{\tau_1, \cdots, \tau_{m-1}, \beta\}$ is a basis of V and, relative to this basis, A has matrix $D(\mu)$.

Assume now that A is a transvection. Since $\mu = 1$, condition (**) shows that A has a characteristic vector outside H if and only if $\gamma = 0$. This implies that $A = E$, contradicting the proviso in the definition of transvection excluding E. Therefore A has no characteristic vectors outside H. Let τ_1 be any vector outside H and define $\tau_2 = A\tau_1 - \tau_1$. Note that $A\tau_1 = \tau_1 + \gamma_0$, where $\gamma_0 \in H$ ($\mu = 1$ since A is a transvection); therefore $\tau_2 = \gamma_0 \in H$. Now $\{\tau_1, \tau_2\}$ is independent, lest τ_1 be characteristic. Extend $\{\tau_1, \tau_2\}$ to a basis $\{\tau_1, \cdots, \tau_m\}$ of V, where $\{\tau_3, \cdots, \tau_m\}$ is chosen from H. Relative to this basis, A has matrix $B_{21}(1)$, for $A\tau_1 = \tau_1 + \tau_2$ while $A\tau_i = \tau_i$ for $i \neq 1$. ∎

Corollary 8.20 *All matrices $B_{ij}(\lambda)$, where $i \neq j$ and $\lambda \neq 0$, are conjugate in $GL(m, K)$.*

Proof It suffices to prove that $B_{ij}(\lambda)$ also represents the transvection A. Choose τ_i outside H and define $\tau_j = \lambda^{-1}(A\tau_i - \tau_i)$; for $l \neq i, j$,

choose $\tau_i \in H$ so that $\{\tau_1, \cdots, \tau_m\}$ is a basis of V. Relative to this basis, A has matrix $B_{ij}(\lambda)$. ∎

Corollary 8.21 *Considered as groups of linear transformations, $GL(m, K)$ is generated by all the dilatations and transvections and $SL(m, K)$ is generated by all transvections.*

Proof Theorem 8.19 allows us to translate Lemma 8.11 from matrices to linear transformations. ∎

Let us look at transvections more closely. If A is a transvection of V fixing a hyperplane H and if $\alpha \notin H$, then

$$A(\alpha) = \alpha + \gamma, \qquad \gamma \in H.$$

If $\beta \in V$, then

$$(***) \qquad \beta = b\alpha + \gamma', \qquad b \in K, \gamma' \in H,$$

and equation (*) in the midst of Lemma 8.18 gives

$$A(\beta) = \beta + b\gamma.$$

The function $f\colon V \to K$ defined by $\beta \to b$ is, from (***), K-linear (it is a **functional** on V) and has kernel H. Thus, A determines a functional f and a vector $\gamma \in H = \ker f$ so that A satisfies the formula

$$A(\beta) = \beta + f(\beta)\gamma, \qquad \text{all } \beta \in V.$$

Conversely, given a functional f and $\gamma \in \ker f$, we may define a function $T_{f,\gamma}\colon V \to V$ by

$$T_{f,\gamma}(\beta) = \beta + f(\beta)\gamma,$$

which is easily seen to be a transvection when $f \not\equiv 0$ and $\gamma \neq 0$. Therefore, every transvection is $T_{f,\gamma}$ for some f and γ.

EXERCISES

**8.15. $T_{f,\gamma} \circ T_{f,\gamma'} = T_{f,\gamma+\gamma'}$.

**8.16. $T_{af,\gamma} = T_{f,a\gamma}$ for all $a \in K$.

**8.17. $T_{f,\gamma} = T_{g,\delta}$ if and only if there is a scalar $a \in K$ with $g = af$ and $\gamma = a\delta$.

**8.18. If $S \in GL(m, K)$, then

$$ST_{f,\gamma}S^{-1} = T_{fS^{-1},S\gamma}.$$

Exercise 8.18 says that a conjugate of a transvection is again a transvection. This is precisely why linear transformations are preferable to matrices, for a conjugate of $B_{ij}(\lambda)$ need not be a matrix of the same form.

Theorem 8.22 *The commutator subgroup of $GL(m, K)$ is $SL(m, K)$ unless $(m, K) = (2, GF(2))$.*

Proof We know that $\det: GL \to K^*$ has kernel SL, so that $GL/SL \cong K^*$; since K^* is abelian, $(GL)' \subset SL$.

For the reverse inclusion, let $\pi: GL \to GL/(GL)'$ be the natural map. Since all transvections are conjugate in GL, by Theorem 8.19, $\pi(T) = \pi(T')$ for any two transvections T and T'; let d denote their common value. If we avoid the exceptional case $(2, GF(2))$, every hyperplane H contains nonzero vectors γ, γ' (not necessarily distinct) such that $\gamma + \gamma'$ is also nonzero. Choose a hyperplane H, nonzero vectors γ, γ', $\gamma + \gamma' \in H$, and a functional f having kernel H. By Exercise 8.15, we have in GL

$$T_{f,\gamma} \circ T_{f,\gamma'} = T_{f,\gamma+\gamma'}.$$

Since γ, γ', $\gamma + \gamma'$ are all nonzero, each of these terms is a transvection (they must not be E). Applying π to this equation gives $d^2 = d$ in K^*, whence $d = 1$. Thus, every transvection is in $\ker \pi = (GL)'$. Since SL is generated by transvections, $SL \subset (GL)'$, as desired.

Finally, $GL(2, 2)$ is a genuine exception:

$$GL(2, 2) = SL(2, 2) \cong S_3,$$

while $(S_3)' = A_3$, a proper subgroup. ∎

To this point, we have allowed $m = \dim V \geq 2$; the assumption $m \geq 3$ enters crucially in the next result.

Theorem 8.23 *If $m \geq 3$, then all transvections are conjugate in $SL(m, K)$.*

REMARK We know transvections are conjugate in GL; we claim they are even conjugate in SL when $m \geq 3$.

Proof Let $T_{f,\gamma}$ and $T_{f',\gamma'}$ be transvections, and let H and H' be the hyperplanes fixed by each. Choose vectors β, $\beta' \in V$ with

$$f(\beta) = 1 = f'(\beta').$$

We claim there is a transformation S in $GL(m, K)$ with

$$S(\gamma) = \gamma', \qquad S(H) = H', \qquad \text{and} \qquad S(\beta) = \beta'.$$

There is a basis of H, one of whose terms is γ, say $\{\gamma, \gamma_2, \cdots, \gamma_{m-1}\}$; similarly, there is a basis $\{\gamma', \gamma'_2, \cdots, \gamma'_{m-1}\}$ of H' containing γ'. Since H and H' are hyperplanes and $\beta \notin H$, $\beta' \notin H'$, then $\{\beta, \gamma, \gamma_2, \cdots, \gamma_{m-1}\}$ and $\{\beta', \gamma', \gamma'_2, \cdots, \gamma'_{m-1}\}$ are bases of V. Define $S: V \to V$ to be the transformation taking the first of these ordered bases onto the second.

If $m \geq 3$, we claim that we can further choose S of determinant 1. Suppose $\det S = d$. Since $m \geq 3$, the first basis of V constructed above contains at least one other vector besides β and γ. Redefine S so that $S(\gamma_{m-1}) = d^{-1}\gamma'_{m-1}$. Relative to the basis $\{\beta, \gamma, \gamma_2, \cdots, \gamma_{m-1}\}$, the new transformation has as its matrix that of the original S with the last column multiplied by d^{-1}. This new transformation thus has determinant 1 as well as the other properties of S.

We claim $ST_{f,\gamma}S^{-1} = T_{f',\gamma'}$, which will complete the proof. By Exercise 8.18, we know $ST_{f,\gamma}S^{-1} = T_{fS^{-1},S\gamma}$. Now $S\gamma = \gamma'$; to see that $fS^{-1} = f'$, it suffices to check whether they agree on a basis of V; they do agree on $\{\beta', \gamma', \gamma'_2, \cdots, \gamma'_{m-1}\}$. ∎

EXERCISE

8.19. Show that $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ are not conjugate in $SL(2, 3)$.

Generalize to any field K in which -1 is not a square.

We need a bit more information about transvections before we can prove the main theorem.

Definition If H is a hyperplane, then $\mathbf{T}(H)$ is the set consisting of the identity together with all transvections fixing H.

($\mathbf{T}(H)$ is the analog of the subgroup Γ that arose in the 2×2 case in Lemma 8.16.)

Lemma 8.24 *Let H be a hyperplane of V.*

(i) *There is a functional f with $H = \ker f$ such that*

$$\mathbf{T}(H) = \{T_{f,\gamma} : \gamma \in H\};$$

(ii) $\mathbf{T}(H)$ *is an abelian subgroup of $SL(m, K)$; in fact,* $\mathbf{T}(H) \cong H$;

(iii) $C_{SL}(\mathbf{T}(H)) = Z_0 \cdot \mathbf{T}(H)$ *(recall that Z_0 is the center of SL).*

Proof (i) Assume that $T_{f,\gamma}$ and $T_{f',\gamma'}$ are in $\mathbf{T}(H)$ and that $f \not\equiv 0$. Let $\{\alpha_1, \cdots, \alpha_m\}$ be a basis of V whose first $m - 1$ terms lie in H. If

$f(\alpha_m) = a$ and $f'(\alpha_m) = a'$, then set $b = a'a^{-1}$; it is easy to check that $f' = bf$. By Exercise 8.16,

$$T_{f',\gamma'} = T_{bf,\gamma'} = T_{f,b\gamma'}.$$

(ii) One verifies quickly that $\mathbf{T}(H)$ is a subgroup of SL and that $T_{f,\gamma} \to \gamma$ is an isomorphism $\mathbf{T}(H) \to H$ (where we choose the same functional f for all T's, by (i)).

(iii) Since $\mathbf{T}(H)$ is abelian, we have $C_{SL}(\mathbf{T}(H)) \supset Z_0 \cdot \mathbf{T}(H)$. For the reverse inclusion, suppose $A \in SL$ and

$$AT_{f,\gamma}A^{-1} = T_{f,\gamma}, \qquad \text{all } T_{f,\gamma} \in \mathbf{T}(H).$$

By Exercise 8.18,

$$T_{fA^{-1},A\gamma} = T_{f,\gamma};$$

by Exercise 8.17, there is a scalar $a \in K$ with $fA^{-1} = af$ and $\gamma = aA\gamma$. Therefore

$$A(\gamma) = a^{-1}\gamma, \qquad \text{all } \gamma \in H$$

and

$$fA = a^{-1}f.$$

If $\{\alpha_1, \cdots, \alpha_m\}$ is a basis of V whose first $m - 1$ terms lie in H, then the last two equations give, for $\beta \in V$,

$$A(\beta) = a^{-1}\beta + \gamma', \qquad \gamma' \in H.$$

This says that A is a scalar multiple of a transvection, as desired. ∎

Theorem 8.25 *The groups $PSL(m, K)$ are simple for all $m \geq 3$ and all fields K.*

Proof We show that if $G \lhd SL$ and $G \not\subset Z_0$, then $G = SL$. By Theorem 8.23, it suffices to prove G contains a transvection. Choose $A \in G$ with $A \notin Z_0$. There is thus a transvection T that does not commute with A:

$$B = (T^{-1}AT)A^{-1} \neq E.$$

Note that $B \in G$ since G is normal. Now

$$B = T^{-1}(ATA^{-1}) = T_1T_2,$$

where T_i is a transvection, $i = 1, 2$. If $T_i = T_{f_i,\gamma_i}$ and $H_i = \ker f_i$, then, for all $\beta \in V$,

$$T_i(\beta) = \beta + f_i(\beta)\gamma_i, \qquad \text{where } \gamma_i \in H_i,\ i = 1, 2.$$

Let W be the subspace of V spanned by γ_1 and γ_2, so that

dim $W \leq 2$. Since dim $V \geq 3$, there is a hyperplane H of V containing W. Now

$$(\#) \qquad B(H) \subset H,$$

for if $\eta \in H$, then

$$B(\eta) = T_1T_2(\eta) = T_2(\eta) + f_1(T_2(\eta))\gamma_1 = \eta + f_2(\eta)\gamma_2 + f_1(T_2(\eta))\gamma_1 \in H + W = H.$$

Next, we claim $H_1 \cap H_2 \neq \{0\}$. If $H_1 = H_2$, this is surely true. If $H_1 \neq H_2$, then $H_1 + H_2 = V$ (hyperplanes are maximal subspaces) and $\dim(H_1 + H_2) = m$. It follows from

$$\dim(H_1 \cap H_2) + \dim(H_1 + H_2) = \dim H_1 + \dim H_2$$

that $\dim(H_1 \cap H_2) = m - 2 \geq 1$, whence $H_1 \cap H_2 \neq \{0\}$. Choose $\zeta \in H_1 \cap H_2$, $\zeta \neq 0$. Then

$$(\#\#) \qquad B(\zeta) = T_1T_2(\zeta) = \zeta.$$

We may assume B is not a transvection (or we are done, since $B \in G$). Therefore $B \notin \mathbf{T}(H)$, which is wholly comprised of transvections. If $B = aS$, where $a \in K$ and $S \in \mathbf{T}(H)$, then (##) says that ζ is a characteristic vector of S. By Theorem 8.19, $\zeta \in H$ and so $S(\zeta) = \zeta$. Thus

$$\zeta = B(\zeta) = aS(\zeta) = a\zeta$$

and $a = 1$; hence $B = S \in \mathbf{T}(H)$, a contradiction. Therefore

$$B \notin Z_0 \cdot \mathbf{T}(H) = C_{SL}(\mathbf{T}(H)),$$

by Lemma 8.24. There is thus a transvection $U \in \mathbf{T}(H)$ with

$$C = U^{-1}B^{-1}UB \neq E.$$

Clearly $C = (U^{-1}B^{-1}U)B \in G$. Also, if $\eta \in H$,

$$C(\eta) = U^{-1}B^{-1}UB(\eta) = U^{-1}B^{-1}(B(\eta)) = \eta$$

since $B(\eta) \in H$, by (#), and $U \in \mathbf{T}(H)$ fixes H. Therefore C fixes H. By Theorem 8.19, C is either a dilatation or a transvection. It is not a dilatation, for det $C = 1$. Hence C is a transvection in G, and the proof is complete. ∎

EXERCISES

****8.20.** Let K be a subfield of F and let A be a square matrix over K. If A is similar over F to a matrix B (whose entries may lie in F), then A^n is scalar if and only if B^n is scalar.

**8.21. Let K be a subfield of F, where K is a finite field having q elements; let $m(x)$ be an irreducible polynomial in $K[x]$ of degree d. If α is a root of $m(x)$ in F, then

$$\alpha^{q^d} = \alpha$$

and the roots of $m(x)$ are

$$\alpha, \alpha^q, \alpha^{q^2}, \cdots, \alpha^{q^{d-1}}.$$

(HINT: Show that $(m(x))^q = m(x^q)$ by using Exercises 8.1 and 8.2.)

Observe that $|PSL(3, 4)| = 20,160 = \frac{1}{2}8!$, so that $PSL(3, 4)$ and A_8 have the same order.

Theorem 8.26 *$PSL(3, 4)$ and A_8 are not isomorphic. Therefore, there are nonisomorphic finite simple groups having the same order.*

Proof A_8 contains (12345)(678), an element of order 15. We shall prove that $PSL(3, 4)$ contains no element of order 15, and this will prove the theorem.

Let A be a 3×3 unimodular matrix over $GF(4)$, and let $c(x)$ be its characteristic polynomial. By Theorem 8.5, there is a field F containing $GF(4)$ in which $c(x)$ is a product of linear factors; thus, F contains all the characteristic roots of A.

Since a conjugate of an element in a group has the same order as the element, we need only examine the canonical form of A. Furthermore, the center of SL is the scalar matrices, so that Exercise 8.20 allows us to examine canonical forms over F instead of over $GF(4)$. This last remark allows us to assume that all characteristic roots of A are in F, so that we may use the Jordan canonical form of A.

There are only three possible Jordan canonical forms for a 3×3 matrix:

(i) a 3×3 Jordan block:

$$B = \begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix};$$

(ii) a direct sum of a 1×1 and a 2×2 Jordan block:

$$B = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 1 & b \end{bmatrix};$$

(iii) a direct sum of three 1×1 Jordan blocks:

$$B = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}.$$

. If $\mu \in GF(4)$ is nonzero, then $\mu^3 = 1$. Therefore, if A^{15} (hence B^{15}) is a scalar matrix, then $A^{45} = E$; hence, $B^{45} = E$. Since F has characteristic 2 and 45 is odd, Exercises 4.49 and 4.50 show that odd powers of the first two canonical forms for B cannot be scalar. Only the diagonal matrix (iii) remains.

Suppose the characteristic polynomial $c(x)$ of B is irreducible over $GF(4)$. By Exercise 8.21, the roots a, b, c of $c(x)$ are given by α, α^4, α^{16}, and $\alpha^{64} = \alpha$. Hence, $\alpha^{63} = 1$. Since $B^{45} = E$, $\alpha^{45} = 1$; therefore, $\alpha^9 = 1$, $B^9 = E$, and B does not represent an element of order 15 in *PSL*.

At the other end of the spectrum is the case in which all characteristic roots of B already lie in $GF(4)$. As we remarked above, this implies that $B^3 = E$, so there is no element of order 15 here.

The final case is $c(x) = (x - a)m(x)$, where $a \in GF(4)$ and $m(x)$ is an irreducible polynomial over $GF(4)$ of degree 2. Exercise 8.21 says that the roots of $m(x)$ are α and α^4; furthermore, it says that $\alpha^{16} = \alpha$, i.e., $\alpha^{15} = 1$; hence, α is a 15th root of unity. Since $GF(16)$ contains every 15th root of unity (i.e., every root of $x^{15} - 1$ over $GF(2)$), we may assume that $F = GF(16)$. Let ρ be a primitive 15th root of unity in F. There is an integer t such that $\alpha = \rho^t$ and $\alpha^4 = \rho^{4t}$. Note that the nonzero elements of $GF(4)$ are ρ^5, ρ^{10}, and $\rho^{15} = 1$. Hence,

$$B = \begin{bmatrix} a & 0 & 0 \\ 0 & \rho^t & 0 \\ 0 & 0 & \rho^{4t} \end{bmatrix},$$

where $a = \rho^5$, ρ^{10}, or ρ^{15}. We claim that for any of these choices of a, B^5 is a scalar matrix, and so cannot represent an element of order 15 in *PSL*. For example, suppose $a = \rho^5$. Since B is unimodular, $\rho^{5+5t} = 1$, i.e., $\rho^{5t} = \rho^{10}$. The diagonal entries of B^5 are ρ^{25}, $\rho^{5t} = \rho^{10}$ and $\rho^{20t} = \rho^{40}$, all of which equal ρ^{10}. This completes the proof of the theorem. ∎

The reader is referred to the books of E. Artin, *Geometric Algebra*, and J. Dieudonné, *Sur les Groupes Classiques*, for discussions of some other families of simple groups.

CHAPTER
9

Infinite Abelian Groups

THE FIRST REDUCTION: TORSION AND TORSION-FREE

A valuable viewpoint in studying an infinite abelian group G is to consider G as an extension of more manageable groups. Of course this reduces the study of G to the study of an extension problem and the study of the more manageable groups. In our first reduction, the more manageable groups are torsion and torsion-free.

Before we begin, it is convenient to agree upon notation and to make some quite formal definitions. First, all groups are abelian and are written additively; second, the trivial group having one element is denoted 0 (instead of $\{0\}$).

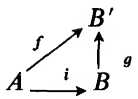
Definition In the following diagram, capital letters denote groups and arrows denote homomorphisms.

$$\begin{array}{ccc} A & \xrightarrow{\alpha} & B \\ {\scriptstyle \beta'}\downarrow & & \downarrow{\scriptstyle \beta} \\ A' & \xrightarrow{\alpha'} & B' \end{array}$$

This diagram **commutes** in case $\beta\alpha = \alpha'\beta'$.

A special case of such a diagram is a triangular diagram, i.e., one of

the homomorphisms is an identity. A common example is



where A is a subgroup of B, and i is the inclusion. This diagram commutes if $gi = f$, i.e., $g|A = f$. Another way of saying this is that g **extends** f.

We say that a larger diagram composed of squares and triangles **commutes** if each component diagram commutes.

Here is a second formal definition.

Definition Let

$$(*) \qquad \cdots \longrightarrow A_{k+1} \xrightarrow{f_{k+1}} A_k \xrightarrow{f_k} A_{k-1} \xrightarrow{f_{k-1}} A_{k-2} \longrightarrow \cdots$$

be a sequence of groups and homomorphisms; this sequence is **exact** in case the image of each map is equal to the kernel of the next map.

EXERCISES

9.1. If $0 \to A \xrightarrow{f} B$ is an exact sequence, then f is one-to-one. (There is no need to label arrows $0 \to A$ or $C \to 0$; any such map must be the zero map.)

9.2. If $B \xrightarrow{g} C \to 0$ is an exact sequence, then g is onto.

9.3. If $0 \to A \to B \to C \to 0$ is an exact sequence, then B is an extension of A by C.

9.4. In the exact sequence (*), f_{k+1} is onto if and only if f_{k-1} is one-to-one.

Definition If G is abelian, the **torsion**[1] **subgroup** of G, denoted tG, is the set of all elements in G of finite order.

Since G is abelian, tG is a subgroup of G.

Definition A group G is **torsion** in case $tG = G$; G is **torsion-free** in case $tG = 0$.

Theorem 9.1 *Every abelian group G is an extension of a torsion group by a torsion-free group.*

[1] This terminology comes from algebraic topology, where a space is "twisted" it if has homology groups containing elements of finite order.

Proof It suffices to prove that G/tG is torsion-free. Suppose $n\bar{x} = 0$ for some $\bar{x} \in G/tG$ and some integer $n \neq 0$. If we lift $\bar{x}$ to $x \in G$, then $nx \in tG$; hence, there is an integer $m \neq 0$ with $mnx = 0$. Since $mn \neq 0$, $x \in tG$, and so $\bar{x} = 0$. ∎

Theorem 9.1 reduces the study of arbitrary abelian groups to the study of torsion groups, torsion-free groups, and an extension problem. Our first question is to determine whether this particular extension problem is only virtual or if there is a group whose torsion subgroup is not a direct summand.[2] Let us first generalize one of our methods of manufacturing groups.

Let K be a nonempty set (possibly infinite), and for each $k \in K$, let there be a given a group A_k.

Definition The **direct product**[3] of the A_k, denoted $\prod_{k\in K} A_k$ is the group consisting of all elements (a_k) in the cartesian product of the A_k under the binary operation:

$$(a_k) + (a_k') = (a_k + a_k').$$

Definition The **direct sum** of the A_k, denoted $\sum_{k\in K} A_k$, is the subgroup of $\prod_{k\in K} A_k$ consisting of all elements (a_k) almost all of whose coordinates are 0, i.e., only finitely many a_k are nonzero.

If the index set K is finite, then $\prod_{k\in K} A_k = \sum_{k\in K} A_k$; if the index set K is infinite, the product and the sum are distinct (and are isomorphic only in rare cases).

EXERCISES

9.5. Let $\{A_k: k \in K\}$ be a family of abelian groups. Prove that $t(\prod A_k) \subset \prod tA_k$ and $t(\sum A_k) = \sum tA_k$. Show that the first inclusion is proper for the special case when the index set K is the positive integers and $A_k = \sigma(p^k)$ for some prime p.

9.6. Define functions $p_i: \prod_{k\in K} A_k \to A_i$ by $p_i((a_k)) = a_i$. Prove that each p_i is a homomorphism of $\prod_{k\in K} A_k$ onto A_i (p_i is called the ith **projection**.) The restriction of p_i (also called a projection) maps $\sum_{k\in K} A_k$ onto A_i as well.

[2] If an abelian group is a semidirect product, then it is a direct product (or, in additive terminology, a direct sum).

[3] Also called the *strong direct sum*, or the *complete direct sum*.

As with finite direct sums, there is an internal version of infinite direct sums.

Definition Let $\{A_k: k \in K\}$ be a family of subgroups of G. Then $G = \sum A_k$ **(internal direct sum)** if $G = \langle \bigcup A_k \rangle$ and, for each i, $A_i \cap \langle \bigcup_{k \neq i} A_k \rangle = 0$.

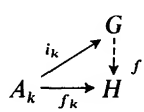
If G is the (external) direct sum $G = \sum A_k$, then G is the internal direct sum of its subgroups $\{A_k': k \in K\}$, where A_k' consists of all those "vectors" having 0 in every coordinate but the kth.

EXERCISES

****9.7.** Let $\{A_k: k \in K\}$ be a family of subgroups of G. Prove that $G = \sum A_k$ (internal) if and only if every nonzero element $g \in G$ has a unique expression of the form $g = a_{k_1} + \cdots + a_{k_n}$, where $a_{k_i} \in A_{k_i}$, the k_i are distinct, and each $a_{k_i} \neq 0$.

****9.8.** Let A be a subgroup of G. Prove that A is a direct summand of G (i.e., there is a subgroup B of G so that G is the internal direct sum of A and B) if and only if there is a homomorphism $p: G \to A$ with $p(a) = a$ for all $a \in A$.

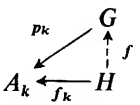
****9.9.** Let $\{A_k: k \in K\}$ be a family of subgroups of G. Prove that $G \cong \sum A_k$ if and only if, given any abelian group H and any set of homomorphisms $f_k: A_k \to H$, there exists a unique homomorphism $f: G \to H$ that extends each f_k; i.e., if $i_k: A_k \to G$ is the inclusion, there is a unique f such that all the following diagrams commute:



(HINT: First consider the case $H = \sum A_k$ in order to get maps $f: G \to \sum A_k$ and $g: \sum A_k \to G$. To show that f is an isomorphism, use the uniqueness hypothesis in a second diagram in which $H = G$.)

****9.10.** Let $\{A_k: k \in K\}$ be a family of subgroups of G. Prove that $G \cong \coprod A_k$ if and only if there exist homomorphisms $p_k: G \to A_k$ with $p_k|A_k$ = identity, and given any abelian group H and any set of homomorphisms $f_k: H \to A_k$, there is a unique homomorphism

$f\colon H \to G$ such that $p_k f = f_k$ for all k, i.e., there is a unique f such that all the following diagrams commute:



Definition Let $x \in G$ and let n be an integer; x is **divisible by** n in case there is an element $y \in G$ with $ny = x$.

EXERCISES

9.11. Let $x \in G$. Any two solutions in G to the equation $ny = x$ differ by an element z with $nz = 0$.

****9.12.** Let $x \in G$ have order n; if $(m, n) = 1$, then x is divisible by m.

Theorem 9.2 *There exists an abelian group G whose torsion subgroup is not a direct summand.*

Proof Let P be the set of all primes and let $G = \prod_{p \in P} \sigma(p)$.

We claim that there is no nonzero element x in G that is divisible by every prime p. If $qy = x$, then $(qy_p) = (x_p)$, i.e., $qy_p = x_p$ for every p. In particular, if $q = p$, then $x_p = 0$. Therefore, if x is divisible by every prime p, then each coordinate of x is 0, and so $x = 0$.

We shall now show that G/tG contains a nonzero element that is divisible by every prime. Were tG a direct summand of G, then $G \cong tG \oplus (G/tG)$ would also contain such a nonzero element, contradicting the first part of the proof. Consider the element $(a_p) + tG$ in G/tG, where a_p is a generator of $\sigma(p)$. If q is a prime, then, by Exercise 9.12, for each $p \neq q$, there is an element $x_p \in \sigma(p)$ with $qx_p = a_p$. Hence, if we define $x_q = 0$, then

$$q(x_p) = (a_p) - y$$

where y has 0 in each coordinate save the qth, where it has a_q. Therefore, $y \in tG$ and

$$q((x_p) + tG) = (a_p) - y + tG = (a_p) + tG.$$

We have shown that $(a_p) + tG$ is divisible by every prime; since this element is nonzero, tG cannot be a direct summand of G. ∎

Theorem 9.3 *Every torsion group G is the direct sum of p-primary groups.*

Proof For any prime p, let

$$G_p = \{x \in G: x \text{ has order some power of } p\}.$$

(G_p is the ***p*-primary component** of G.) The reader may now prove that $G = \sum G_p$, using our proof of the primary decomposition theorem (Theorem 4.4) as his model. ∎

Theorem 9.4 *Let G and H be torsion; $G \cong H$ if and only if $G_p \cong H_p$ for all primes p.*

Proof Let $f: G \to H$ be an isomorphism and let $g: H \to G$ be its inverse. One checks easily that $f(G_p) \subset H_p$ and, by symmetry, that $g(H_p) \subset G_p$. Let $f_p = f|G_p$ and $g_p = g|H_p$. Both $f_p g_p$ and $g_p f_p$ are identities, so that $G_p \cong H_p$.

Conversely, if $f_p: G_p \to H_p$ are isomorphisms, then there is an isomorphism $f: G \to H$ defined by $(x_p) \to (f_p(x_p))$. ∎

Because of these two theorems, the study of torsion groups is reduced to the study of p-primary groups.

THE SECOND REDUCTION: DIVISIBLE AND REDUCED

Had we asked the reader in Chapter 1 to give examples of infinite abelian groups, he probably would have responded with the integers, the rationals, and the reals. In this section, we study a common generalization of the latter two groups, the divisible groups. We shall see that every group is an extension of a divisible group by a group having no divisible subgroups.

Definition A group G is **divisible** in case each $x \in G$ is divisible by every $n > 0$.

EXERCISES

9.13. We shall henceforth denote the additive group of rationals by $\mathbf{Q}$. Prove that the following groups are divisible: $\mathbf{Q}$; additive group of real numbers $\mathbf{R}$; additive group of complex numbers $\mathbf{C}$; multiplicative group of positive reals; multiplicative group of

nonzero elements of an algebraically closed field; the circle group **T**, i.e., the multiplicative group of all complex numbers z with $|z| = 1$.

*9.14. The group G/tG constructed in Theorem 9.2 is divisible.

**9.15. A quotient of a divisible group is divisible.

9.16. A direct sum (direct product) of groups is divisible if and only if each summand (factor) is divisible.

9.17. A torsion-free divisible group is a vector space over **Q. (HINT: Verify the axioms after noting that a solution to $ny = x$ exists and is unique.)

The reader knows that every finite-dimensional vector space V over a field F has a basis and that any two bases of V have the same number of elements. We now prove the infinite analogs of these theorems; because this may be the reader's first contact with Zorn's lemma (see Appendix IV), we proceed leisurely.

Definition Let V be a vector space over a field F. A subset X of V is **dependent** in case there exist a finite number of vectors $x_1, x_2, \cdots, x_n$ in X and nonzero scalars $a_1, a_2, \cdots, a_n \in F$ such that $\sum a_i x_i = 0$; otherwise, X is **independent**. A subset X **spans** V in case every vector in V is a finite linear combination of vectors in X. A **basis** of V is an independent subset that spans V.

Example Let $V = \mathbf{Q}[x]$, the vector space of all polynomials in x over **Q**. A basis of V is the set $\{1, x, x^2, \cdots, x^m, \cdots\}$.

Theorem 9.5 *If V is a vector space over a field F, then V has a basis. In fact, every independent subset I of V is contained in a basis.*

Proof Let $\mathscr{S}$ be the family of all independent subsets of V containing I; partially order $\mathscr{S}$ by ordinary inclusion. Let $\{X_\alpha : \alpha \in A\}$ be a simply ordered subset of $\mathscr{S}$, i.e., the X_α are independent subsets of V containing I, and given any two of them, one contains the other. From this it follows that, given any finite number of these X_α, one contains all the others. Now let X be the union of these X_α. It is trivial that X contains each X_α, but we must verify that $X \in \mathscr{S}$ in order that it be an upper bound. Suppose $x_1, x_2, \cdots, x_n \in X$ and $\sum a_i x_i = 0$. Now each x_i got into X by being in X_{α_i} for some α_i. There being only finitely many X_{α_i}, one contains all the others; hence, $x_1, x_2, \cdots, x_n$ all lie in this one X_α, which is, by hypothesis, independent. Therefore, each $a_i = 0$ and X is independent.

By Zorn's lemma, there is a maximal independent subset Y of V that contains I. We claim that Y is a basis of V, for which it now suffices to prove that Y spans V. Suppose $x \in V$ and x is not a linear combination of elements in Y. Consider the set $Y' = Y \cup \{x\}$. Y' is an independent set containing I, as the reader may easily check, and this contradicts the maximality of Y. ∎

Corollary 9.6 *Every subspace W of a vector space V is a direct summand of V.*

Proof Let X be a basis of W. Since X is independent, Theorem 9.5 says, there exists a subset X' of V such that $X \cup X'$ is a basis of V. If W' is the subspace spanned by X', then it is easy to see that $V = W \oplus W'$. ∎

Theorem 9.7 *Let V be a vector space over F. Considering V as an abelian group, V is a direct sum of copies of F.*

Proof Let $B = \{x_k : k \in K\}$ be a basis of V and let F_k denote the one-dimensional subspace generated by x_k. Clearly, each F_k is isomorphic, as a group, to the additive group F.

We claim that the additive group V is isomorphic to $\sum_{k \in K} F_k$. Since B spans V, every nonzero vector $\alpha \in V$ has an expression $\alpha = \sum r_{k_i} x_{k_i}$, where the r are nonzero elements of F and all the x are distinct; furthermore, each $r_{k_i} x_{k_i} \in F_{k_i}$. Since B is independent, this is the only expression for α of this kind. By Exercise 9.7, $V \cong \sum F_{k_i}$. ∎

Corollary 9.8 (i) *Every torsion-free divisible group G is a direct sum of copies of* $\mathbf{Q}$;

(ii) *An abelian group G in which every nonzero element has prime order p is a direct sum of copies of $\sigma(p)$.*

Proof (i) By Exercise 9.17, G is a vector space over $\mathbf{Q}$. (ii) In the proof of Lemma 4.5 it was verified that G admits a scalar multiplication by elements in Z_p. ∎

We see in particular that the additive group of real numbers is a vector space over $\mathbf{Q}$. A basis of this vector space is usually called a **Hamel basis**, and it is useful for constructing certain analytical counterexamples. For example, a Hamel basis may be used to exhibit a discontinuous function f on the reals satisfying the functional equation

$$f(x + y) = f(x) + f(y)$$

for all real numbers x and y. (If c is the cardinal of the continuum, there are 2^c such functions f but only c continuous real-valued functions.)

EXERCISES

9.18. Let G be a (not necessarily abelian) group, and suppose $|\mathrm{Aut}(G)| = 1$. Prove that G has at most two elements. (The reader now has the power to complete Exercise 7.10.)

****9.19.** Let V be a vector space over a field F. Prove that any two bases of V have the same number of elements.
(HINTS: (1) The reader need only consider the case in which V is infinite dimensional, for the finite-dimensional case is well known. (2) The following theorem of set theory may be used: If X is an infinite set and S is the collection of all finite subsets of X, then X and S have the same number of elements.)

Definition Let V_1 and V_2 be vector spaces over F. V_1 and V_2 **have the same dimension** if there are bases B_1 of V_1 and B_2 of V_2 that have the same number of elements.

EXERCISE

9.20. Let V and W be vector spaces over F, where F is isomorphic to either $\mathbf{Q}$ or Z_p. Show that, as abelian groups, $V \cong W$ if and only if V and W have the same dimension. (This exercise is false for arbitrary fields F. For example, a 1-dimensional real vector space is isomorphic, as an abelian group, to a 2-dimensional real vector space.)

The easiest example to exhibit of a torsion divisible group is $\mathbf{Q}/Z$; in particular, its p-primary components are p-primary divisible groups.

Definition If p is a prime, $\sigma(p^\infty)$ denotes the p-primary component of $\mathbf{Q}/Z$.

EXERCISES

9.21. Let $A^{(p)}$ denote the set of all rationals between 0 and 1 of the form m/p^n, where $m, n \geq 0$, under the binary operation "addition modulo 1". For example, if $p = 2$, then $\frac{1}{2} + \frac{1}{2} = 0$, $\frac{1}{2} + \frac{3}{4} = \frac{1}{4}$,

etc. Prove that $A^{(p)}$ is a p-primary group and that $\mathbf{Q}/Z \cong \sum A^{(p)}$; conclude that $A^{(p)} \cong \sigma(p^\infty)$.

9.22. Show that $\sigma(p^\infty)$ is generated by elements $a_1, a_2, a_3, \cdots$, where $pa_1 = 0$, $pa_2 = a_1, \cdots, pa_{n+1} = a_n, \cdots$. If $\langle a_n \rangle$ is the cyclic subgroup of $\sigma(p^\infty)$ generated by a_n, then $\langle a_n \rangle \cong \sigma(p^n)$, $\langle a_n \rangle \subset \langle a_{n+1} \rangle$ for all n, and

$$\sigma(p^\infty) = \bigcup_{n=1}^{\infty} \langle a_n \rangle.$$

9.23. $\sigma(p^\infty)$ is a p-group containing a unique subgroup of order p. Conclude that Theorem 6.31 is false without assuming the group is finite.

9.24. For each $n > 0$, $\mathbf{Q}/Z$ contains a unique subgroup of order n (which must be cyclic).

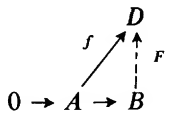
9.25. Every proper subgroup of $\sigma(p^\infty)$ is finite, and the set of subgroups is well ordered by inclusion. (It is an open question whether there exists an infinite nonabelian group all of whose proper subgroups are finite.)

9.26. Prove that $\sigma(p^\infty)$ has the DCC but not the ACC.

9.27. For a fixed prime p, let G be the set of all pth power roots of unity, i.e., all complex numbers of the form $\exp(2\pi ik/p^n)$, where $k \in Z$ and $n \geq 0$. Prove that G is a multiplicative group isomorphic to $\sigma(p^\infty)$.

Our immediate goal is the classification of divisible groups.

Theorem 9.9 (The Injective Property) *Let A be a subgroup of B and let $f\colon A \to D$ be a homomorphism, where D is divisible. Then f can be extended to a homomorphism $F\colon B \to D$, i.e., an F exists making the adjoined diagram commute.*



Proof We use Zorn's lemma. Consider the set $\mathscr{S}$ of all pairs (S, h), where S is a subgroup of B containing A and $h\colon S \to D$ extends f; $\mathscr{S}$ is nonempty, for $(A, f) \in \mathscr{S}$. We partially order $\mathscr{S}$ be decreeing $(S_1, h_1) \leq (S_2, h_2)$ in case $S_1 \subset S_2$ and h_2 extends h_1. If $\{(S_\alpha, h_\alpha)\}$ is a simply ordered subset of $\mathscr{S}$, define (S_0, h_0) as follows: $S_0 = \bigcup_\alpha A_\alpha$; if $s \in S_0$, then $s \in S_\alpha$ for some α; define $h_0(s) = h_\alpha(s)$. We

leave to the reader the proof that $(S_0, h_0) \in \mathscr{S}$ and that it is an upper bound of $\{(S_\alpha, h_\alpha)\}$. By Zorn's lemma, there exists a maximal pair, (M, h). We shall show that $M = B$, which will complete the proof.

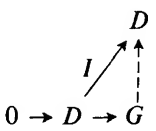
Suppose there is an element $b \in B$ that is not in M. Define $M_1 = M + \langle b \rangle$; clearly, $M \subsetneq M_1$, so that it suffices to extend h to M_1 to reach a contradiction.

CASE (i) $M \cap \langle b \rangle = 0$. Then $M_1 = M \oplus \langle b \rangle$. Define $g: \langle b \rangle \to D$ to be the zero map. By Exercise 9.8, there is a map $F: M_1 \to D$ extending h (and g).

CASE (ii) $M \cap \langle b \rangle \neq 0$. Let k be the smallest positive integer for which $kb \in M$; then every element y in M_1 has the unique expression $y = m + tb$, where $0 \leq t < k$. Since D is divisible, there is an element $x \in D$ with $kx = h(kb)$ (which is defined because $kb \in M$). Define $F: M_1 \to D$ by $F(m + tb) = h(m) + tx$. We leave the straightforward computation that F is a homomorphism extending h to the reader. ∎

Corollary 9.10 *Let D be a subgroup of G, where D is divisible. Then D is a direct summand of G.*

Proof Consider the diagram



where I is the identity map. By Theorem 9.9, there is a homomorphism $p: G \to D$ such that $p(d) = d$ for every $d \in D$. By Exercise 9.8, D is a direct summand of G. ∎

EXERCISES

9.28. An abelian group G is divisible if and only if it has the injective property. (HINT: Extend homomorphisms from nZ into G to homomorphisms from Z into G.)

****9.29.** G is divisible if and only if $pG = G$ for every prime p; if G is p-primary, G is divisible if and only if $G = pG$.

***9.30.** Let G be a nonzero group. Prove that there exists a nontrivial homomorphism $f: G \to \mathbf{Q}/Z$.

9.31. A group G is divisible if and only if every nonzero quotient of G is infinite.

9.32. A group G is divisible if and only if G has no maximal subgroups.

Definition If G is an abelian group, dG is the subgroup of G generated by all the divisible subgroups of G.

Lemma 9.11 *dG is a divisible subgroup of G (which contains every divisible subgroup of G).*

Proof Since the parenthetical remark follows immediately from the definition of dG, we need prove only that dG is divisible. Let $n > 0$ and let $x \in dG$; then $x = x_1 + \cdots + x_k$, where each x_i is in a divisible subgroup D_i of G. Since D_i is divisible, there is an element $y_i \in D_i$ with $ny_i = x_i$. Hence, $y_1 + \cdots + y_k \in dG$ and $n(y_1 + \cdots + y_k) = x$. ∎

Definition An abelian group G is **reduced** if $dG = 0$.

Theorem 9.12 *Every abelian group $G = dG \oplus R$, where R is reduced.*

Proof Since dG is divisible, $G = dG \oplus R$ for some subgroup R (by Corollary 9.10). If R contains a divisible group D, then $dG \oplus D$ is a divisible subgroup of G. But dG contains every divisible subgroup of G, so $D = 0$ and R is reduced. ∎

EXERCISE

9.33. Let G and H be abelian groups. $G \cong H$ if and only if $dG \cong dH$ and $G/dG \cong H/dH$.

The reader should note the similarity of the roles of the subgroups tG and dG. We have seen that every abelian group is an extension (not necessarily split) of a torsion group by a torsion-free group. Now we see that every abelian group is also an extension (always split) of a divisible group by a reduced group.

Lemma 9.13 *Let G and H be divisible p-primary groups. Then $G \cong H$ if and only if $G[p] \cong H[p]$.*[4]

Proof Necessity is simple, so we need prove only sufficiency. Let $f: G[p] \to H[p]$ be an isomorphism. We may consider f as mapping $G[p] \to H$, so that the injective property implies the existence of a map $F: G \to H$ extending f; we claim that F is an isomorphism.

[4] Recall that $G[p] = \{x \in G: px = 0\}$.

(i) F is one-to-one. Let x be a nonzero element of G of order p^n; we show by induction on n that if $F(x) = 0$, then $x = 0$. If $n = 1$, then $x \in G[p]$ and $F(x) = f(x)$; since f is an isomorphism, $x = 0$. Suppose now that x has order p^{n+1} and $F(x) = 0$. If we set $y = px$, then y has order p^n and $F(y) = 0$. By induction, $y = 0$, which contradicts the fact that x has order p^{n+1}.

(ii) F is onto. Let y be a nonzero element of H of order p^n; we show by induction on n that y is in the image of F. If $n = 1$, then $y \in H[p]$ so that $y \in \text{image } f \subset \text{image } F$. Suppose $p^{n+1}y = 0$ and $p^n y \neq 0$. By induction, there is an $x \in G$ with $F(x) = p^n y$. Since G is divisible, there is a $z \in G$ with $p^n z = x$. Thus, $p^n(y - F(z)) = 0$. Using induction again, there is an element $z' \in G$ with $F(z') = y - F(z)$, and so $F(z' + z) = y$. ∎

Theorem 9.14 *Every divisible group D is a direct sum of copies of* $\mathbf{Q}$ *and of copies of* $\sigma(p^\infty)$ *(for various primes p).*

Proof It is easy to check that tD is divisible, so we may assume $D \cong tD \oplus (D/tD)$. Since D/tD is a torsion-free divisible group, it is a direct sum of copies of $\mathbf{Q}$, by Corollary 9.8. Let G be the p-primary component of tD; G is divisible. If dimension $G[p]$ (as a vector space over the integers modulo p) is r, let H be the direct sum of r copies of $\sigma(p^\infty)$. Now H is a p-primary divisible group with $H[p] \cong G[p]$. By Lemma 9.13, $G \cong H$. ∎

The structure of divisible groups is not very complicated. One question yet remains: When are two divisible groups isomorphic? If D is a divisible group, let $D_\infty = D/tD$ and $D_p = (tD)[p]$; observe that D_∞ is a vector space over $\mathbf{Q}$ and D_p is a vector space over Z_p.

Theorem 9.15 *If D and D' are divisible groups, then $D \cong D'$ if and only if* (i) *D_∞ and D'_∞ have the same dimension*; (ii) *for each p, D_p and D'_p have the same dimension.*

Proof Left to the reader. ∎

EXERCISES

9.34. Let G and H be divisible groups, each of which is isomorphic to a subgroup of the other. Prove that $G \cong H$. Is this true if we drop the adjective "divisible"?

9.35. If G and H are divisible and $G \oplus G \cong H \oplus H$, prove that $G \cong H$.

9.36. Prove that the following groups are all isomorphic: $\mathbf{R}/Z$; the circle group $\mathbf{T}$; $\prod_p \sigma(p^\infty)$; $\mathbf{Q}/Z \oplus \mathbf{R}$.

The reader who knows enough field theory is now able to describe the multiplicative group of the algebraic closure of Z_p (it is $\sum_{q \neq p} \sigma(q^\infty)$).

****9.37.** Let $0 \to A \to B \to C \to 0$ be an exact sequence. If A and C are reduced, then B is reduced.

FREE ABELIAN GROUPS

We now consider an important class of groups that is, in a certain sense, dual to the divisible groups. The reader will see properties of these groups that will remind him of several of the theorems we have just proved.

Definition F is a **free abelian group** on $\{x_k : k \in K\}$ in case F is a direct sum of infinite cyclic groups Z_k, where $Z_k = \langle x_k \rangle$.

Theorem 9.16 *If F is free on $\{x_k : k \in K\}$, every nonzero element $x \in F$ has the unique expression*

$$x = m_{k_1} x_{k_1} + \cdots + m_{k_n} x_{k_n},$$

where the m are nonzero integers and the k_i are distinct.

Proof Exercise 9.7. ∎

Theorem 9.17 *Let*

$$F = \sum_{i \in I} Z_i \qquad \text{and} \qquad G = \sum_{j \in J} Z_j$$

be free abelian groups. Then $F \cong G$ if and only if I and J have the same number of elements.

Proof Suppose F is free on $\{x_i : i \in I\}$. If p is a prime, then F/pF is a vector space over Z_p. For $a \in F$, set $\bar{a} = a + pF$. We claim that $\{\bar{x}_i : i \in I\}$ is a basis of F/pF. It is clear they span; let us prove the set of them is independent.

Let $\sum \bar{m}_i \bar{x}_i = 0$, where $\bar{m}_i \in Z_p$ and not all $\bar{m}_i = \bar{0}$. If m_i is a representative of $\bar{m}_i$ with $0 \leq m_i < p$, then $\sum m_i \bar{x}_i = 0$. In F this equation becomes $\sum m_i x_i \in pF$, i.e.,

$$\sum m_i x_i = p \sum n_i x_i.$$

By Theorem 9.16, $m_i = pn_i$ for all i, so that $\overline{m}_i = \overline{0}$ for all i. This contradiction proves independence. Exercise 9.19 gives $|I| =$ dimension F/pF. Therefore, if $F \cong G$, then $|I| = \dim F/pF = |J|$.

Sufficiency is easy and is left to the reader. ∎

Definition Let F be free on $\{x_i : i \in I\}$ and let G be free on $\{y_j : j \in J\}$; F and G **have the same rank** if I and J have the same number of elements. If I is finite and has n elements, we say that F has **rank** n.

Theorem 9.17 says that two free abelian groups F and G are isomorphic if and only if they have the same rank. In particular, if $F = G$, any two free sets of generators of F have the same number of elements. The reader will not be misled by the analogy: vector space—free abelian group; basis—free set of generators; dimension—rank. In order to stress this analogy, we make the following definition.

Definition A **basis** of a free abelian group F is a free set of generators of F.

Theorem 9.18 *Let F be free with basis $\{x_k : k \in K\}$, G an arbitrary abelian group, and $f : \{x_k : k \in K\} \to G$ any function. There is a unique homomorphism $g : F \to G$ such that*

$$g(x_k) = f(x_k) \qquad \text{for all } k.$$

Proof If $Z_k = \langle x_k \rangle$, we define $f_k : Z_k \to G$ by $f_k(mx_k) = mf(x_k)$. It is clear that each f_k is a homomorphism, so that the hypotheses of Exercise 9.9 are satisfied. ∎

Corollary 9.19 *Every abelian group G is a quotient of a free abelian group.*

Proof We first show that if X is any set, there exists a free abelian group F having X as a basis. If X consists of one element x, an infinite cyclic group Zx that has x as a generator can be constructed. For the general case, set $F = \sum_{x \in X} Zx$.

To prove the corollary, let F be the free abelian group with basis G. By Theorem 9.18, the identity function on G extends to a homomorphism $g : F \to G$, and g is clearly onto. Therefore G is a quotient of F. ∎

The construction of a free abelian group on an arbitrary set is quite convenient. For example, in algebraic topology, one wishes to add and

subtract continuous functions from an n-simplex to a topological space. This is done by forming the free abelian group on all such functions.

Theorem 9.20 (The Projective Property) *Let $\beta: B \to C$ be a homomorphism of B onto C. If F is free and $\alpha: F \to C$ is a homomorphism, then there is a homomorphism $\gamma: F \to B$ with $\beta\gamma = \alpha$, i.e., there is a γ making the diagram below commute.*

$$\begin{array}{ccccc} & & F & & \\ & \swarrow^{\gamma} & \downarrow{\scriptstyle \alpha} & & \\ B & \xrightarrow[\beta]{} & C & \to & 0 \end{array}$$

Proof Let $X = \{x_k: k \in K\}$ be a basis of F. Since β is onto, for each k there is an element $b_k \in B$ with $\beta(b_k) = \alpha(x_k)$. Define a function $f: X \to B$ by $f(x_k) = b_k$. By Theorem 9.18, there is a homomorphism $\gamma: F \to B$ such that $\gamma(x_k) = b_k$. In order to check that $\beta\gamma = \alpha$, it suffices to evaluate each on a set of generators of F, e.g., on X. But $\beta\gamma(x_k) = \beta(b_k) = \alpha(x_k)$ as desired. ∎

Corollary 9.21 *Let G be an abelian group and let $\beta: G \to F$ be onto, where F is free. Then*

$$G = \text{kernel } \beta \oplus S,$$

where $S \cong F$.

Proof Consider the diagram

$$\begin{array}{ccccc} & & F & & \\ & \swarrow^{\gamma} & \downarrow{\scriptstyle I} & & \\ G & \xrightarrow[\beta]{} & F & \to & 0, \end{array}$$

where I is the identity map. Since F has the projective property, there is a homomorphism $\gamma: F \to G$ with $\beta\gamma = I$. By set theory, γ is one-to-one, so $S = \text{image } \gamma \cong F$. We let the reader prove, using Theorem 4.1, that $G = \text{kernel } \beta \oplus S$. ∎

Another way of stating Corollary 9.21 is that G/K free implies K is a direct summand of G.

Theorem 9.22 *Every subgroup H of a free abelian group F is free; moreover, rank $H \leq$ rank F.*

Proof We give two proofs; the first proof works only when F has finite rank, but it allows us to focus on essentials.

Suppose F has finite rank n; we do an induction on n. If $n = 1$, then $F \cong Z$ and the division algorithm shows any subgroup H of F is 0 or isomorphic to Z. Thus H is free and rank $H \leq 1 =$ rank F. For the inductive step, let $\{x_1, \cdots, x_n\}$ be a basis of F, $F_n = \langle x_1, \cdots, x_{n-1}\rangle$, and $H_n = H \cap F_n$. By induction, H_n is free of rank $\leq n - 1$. Now

$$H/H_n = H/H \cap F_n \cong H + F_n/F_n \subset F/F_n \cong Z.$$

By Corollary 9.21, $H = H_n$ or $H = H_n \oplus \langle h\rangle$, where $\langle h\rangle \cong Z$. Therefore H is free of rank $\leq n$.

We now give a second proof that does not assume the rank of F is finite. Let $\{x_k : k \in K\}$ be a basis of F which we assume is well-ordered (that every nonempty set can be somehow well-ordered may be taken as an axiom of set theory in place of the axiom of choice).

For each k, define $F_k = \sum_{j<k} \langle x_j\rangle$ and $H_k = H \cap F_k$. Thus $F = \bigcup F_k$ and $H = \bigcup H_k$; also, $H_k = H \cap F_k = H_{k+1} \cap F_k$. Hence

$$H_{k+1}/H_k = H_{k+1}/(H_{k+1} \cap F_k) \cong (H_{k+1} + F_k)/F_k \subset F_{k+1}/F_k \cong Z.$$

By Corollary 9.21, either $H_{k+1} = H_k$ or

$$H_{k+1} = H_k \oplus \langle h_k\rangle, \qquad \text{where } \langle h_k\rangle \cong Z.$$

We claim H is free on the set of h_k's; note that it will then follow that rank $H \leq$ rank F, for the set of h_k clearly has cardinality $\leq |K| =$ rank F.

Let H^0 be the subgroup of H generated by the h_k. Since $F = \bigcup F_k$, each $h \in H$ (as any element of F) lies in some F_{k+1}. Let $\mu(h)$ be the least index k with $h \in F_{k+1}$. Suppose $H \neq H^0$ and consider $\{\mu(h) : h \in H \text{ and } h \notin H^0\}$. There is a least such index j, for K is well-ordered. Choose $h' \in H$ with $\mu(h') = j$ and $h' \notin H^0$. Now $\mu(h') = j$ says $h' \in H \cap F_{j+1}$, so that

$$h' = a + mh_j, \qquad a \in H_j, m \in Z.$$

Therefore $a = h' - mh_j \in H$, $a \notin H^0$ (lest $h' \in H^0$), and $\mu(a) < j$, a contradiction. Hence $H = H^0$.

Next, we show that linear combinations of the h_k are unique. It suffices to show that if

$$m_1 h_{k_1} + \cdots + m_n h_{k_n} = 0, \qquad k_1 < \cdots < k_n,$$

then each $m_i = 0$. We may assume $m_n \neq 0$. But then $m_n h_{k_n} \in \langle h_{k_n} \rangle \cap H_{k_n} = 0$, a contradiction. This shows that H is free on the h_k. ∎

EXERCISES

9.38. An abelian group is finitely generated if and only if it is a quotient of a free abelian group of finite rank. Conclude that a direct summand of a finitely generated abelian group is also finitely generated.

9.39. Every subgroup H of a finitely generated abelian group G is itself finitely generated; if G can be generated by r elements, so can H. (We shall show later that both statements are false if we delete the adjective "abelian".)

9.40. The multiplicative group of positive rationals is a free abelian group of (countably) infinite rank. (See Exercise 1.26.)

9.41. If F is a free abelian group of finite rank n, then $\mathrm{Aut}(F)$ is isomorphic to the multiplicative group of all $n \times n$ matrices of determinant ± 1 that have entries in Z.

***9.42.** Let $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$ be an exact sequence of free abelian groups. Prove that $B \cong \mathrm{im}\, f \oplus \ker h$. (This exercise is useful in algebraic topology.)

****9.43.** Let F be a free abelian group of rank n and let H be a subgroup of rank $k < n$. Prove that F/H contains an element of infinite order.

***9.44.** An abelian group is free if and only if it has the projective property.

9.45. Let $0 \to F_1 \to F_2 \to \cdots \to F_n \to 0$ be an exact sequence of finitely generated free abelian groups. Prove that

$$\sum_{i=1}^{n} (-1)^i \operatorname{rank} F_i = 0.$$

Theorem 9.23 *Every abelian group G can be imbedded in a divisible group.*

Proof By Corollary 9.19, there is a free abelian group F with $G \cong F/R$ for some subgroup R of F. Now $F = \sum Z$, so that $F \subset \sum \mathbf{Q}$ (just imbed each Z in a copy of the rationals, $\mathbf{Q}$). Therefore,

$$G \cong F/R \subset (\sum \mathbf{Q})/R,$$

and this last group is divisible, being a quotient of a divisible group. ▮

Corollary 9.24 *An abelian group G is divisible if and only if it is a direct summand of every group containing it.*

Proof Necessity is Corollary 9.10. In order to prove sufficiency, first imbed G in a divisible group D, and then recall that every direct summand of a divisible group is divisible. ▮

There is an analogy between theorems about free abelian groups and divisible groups, which may be formalized as follows: Given a commutative diagram containing exact sequences, its **dual** is the commutative diagram containing exact sequences obtained by reversing the direction of all arrows. For example, the dual of "$0 \to A \to B$" is "$B \to A \to 0$", i.e., subgroup and quotient group are dual. We let the reader prove that "direct summand" is its own dual, sum and product are dual (see Exercises 9.9 and 9.10), and projective and injective are dual.

FINITELY GENERATED ABELIAN GROUPS

In this section we apply our techniques to classify an important class of abelian groups.

Theorem 9.25 *Every finitely generated torsion-free abelian group G is free.*

Proof We prove the theorem by induction on n, where $G = \langle x_1, \cdots, x_n \rangle$. If $n = 1$, then G is infinite cyclic (or 0 if $x_1 = 0$) and we are done.

Define $\langle x_n \rangle_* = \{y \in G: my \in \langle x_n \rangle \text{ for some } m \neq 0\}$. It is easy to check that $\langle x_n \rangle_*$ is a subgroup of G and that $G/\langle x_n \rangle_*$ is torsion-free, for $\langle x_n \rangle_*$ is just the inverse image of $t(G/\langle x_n \rangle)$. By induction, $G/\langle x_n \rangle_*$ is free, so that $G = \langle x_n \rangle_* \oplus (\text{free})$. We need only show $\langle x_n \rangle_* \cong Z$ to complete the proof.

If $y \in \langle x_n \rangle_*$, then $my = kx_n$ for some $m \neq 0$; the reader may verify that the map $\langle x_n \rangle_* \to \mathbf{Q}$ given by $y \to k/m$ is a well-defined homomorphism that is one-to-one. Thus, $\langle x_n \rangle_*$ is isomorphic to some finitely generated subgroup of the rationals, say, H. Let $H =$

$\langle a_1/b_1, \cdots, a_t/b_t \rangle$ and let $b = \prod b_i$. Then the map $f: H \to Z$ defined by $h \to bh$ is one-to-one, so that H, hence $\langle x_n \rangle_*$, is isomorphic to a subgroup of Z, and hence is cyclic. ∎

Theorem 9.26 (Basis Theorem) *Every finitely generated abelian group G is a direct sum of cyclic groups.*

Proof G/tG is a finitely generated torsion-free group, and hence is free, by Theorem 9.25. By Lemma 9.21,

$$G \cong tG \oplus (\text{free}).$$

Now tG is finitely generated, being a quotient of the finitely generated group G, so that it is finite (Exercise 4.42). Therefore, tG is a direct sum of cyclic groups, by the basis theorem for finite groups. ∎

Theorem 9.27 (Fundamental Theorem of Finitely Generated Abelian Groups) *Every finitely generated abelian group G is a direct sum of primary and infinite cyclic groups, and the number of summands of each kind depends only on G.*

Proof We know that $G \cong tG \oplus (G/tG)$. The uniqueness for tG is precisely the fundamental theorem for finite abelian groups; the uniqueness of the number of infinite cyclic summands is Theorem 9.17. ∎

We must confess that there are shorter proofs of Theorem 9.26, but this proof is quite natural once one decides to work by induction.

EXERCISES

****9.46.** Prove there is a countable abelian group G that contains an isomorphic copy of every countable abelian group as a subgroup. (HINT: Let G be the direct sum of countably many copies of $\mathbf{Q} \oplus \mathbf{Q}/Z$.)

9.47. Let F be free abelian of finite rank n and let H be a subgroup of rank n; then F/H is finite.

9.48. (Simultaneous Bases) Let F be free abelian of finite rank n and let H be a subgroup of rank n. Then there is a basis $\{x_1, \cdots, x_n\}$ of F and integers $k_1, \cdots, k_n$ such that $\{k_1x_1, \cdots, k_nx_n\}$ is a basis of H. (HINT: Use the fundamental theorem.)

9.49. Let F be free abelian of rank n, and H a subgroup of the same rank; let $\{x_1, \cdots, x_n\}$ be a basis of F, and $\{y_1, \cdots, y_n\}$ a basis of H, so

that $y_i = \sum m_{ij}x_j$. Prove that $[F:H] = |\det(m_{ij})|$. (HINT: First show $|\det(m_{ij})|$ is independent of choice of bases of F and H.)

TORSION GROUPS

Torsion groups can be quite complicated, but there are two special classes of torsion groups that are quite manageable: divisible torsion groups and direct sums of cyclic groups. We shall prove that every torsion group is an extension of a direct sum of cyclics by a divisible group, but our proof requires an investigation of a distinguished kind of subgroup whose definition generalizes that of direct summand. Inasmuch as the definition is technical, it shall be followed by many exercises that will familiarize the reader with the concept.

Definition A subgroup S of G is **pure** in G in case

$$nG \cap S = nS \qquad \text{for every integer } n.$$

It is always true that $nG \cap S \supset nS$, so that it is only the reverse inclusion that is significant. Therefore, purity says that whenever $s = ng$ (where $s \in S$ and $g \in G$), there is an element $s' \in S$ with $s = ns'$. In other words, if an element of S is divisible by n in the big group, it is also divisible by n in the subgroup S.

If $G = \langle x \rangle$ is infinite cyclic and if $S = \langle 2x \rangle$, then S is not pure, for $s = 2x$ is divisible by 2 in G, but there is no element y *in* S with $2y = 2x$.

EXERCISES

9.50. Any direct summand of G is pure in G.

9.51. If G/S is torsion-free, then S is pure; conclude that tG is a pure subgroup of G. Conclude further that a pure subgroup need not be a direct summand.

9.52. If G is torsion-free, a subgroup S of G is pure if and only if G/S is torsion-free.

9.53. Purity is transitive, i.e., if K is pure in H and H is pure in G, then K is pure in G.

9.54. If G is torsion-free, any intersection of pure subgroups is pure. Conclude that if X is a subset of G (where G is torsion-free), then there is a smallest pure subgroup of G containing X; we denote

this subgroup $\langle X \rangle_*$ and call it the **pure subgroup** of G **generated** by X.

9.55. Let $x \in G$, where G is torsion-free. Show that the pure subgroup generated by x is

$$\{y \in G: my \in \langle x \rangle \text{ for some } m \neq 0\}.$$

9.56. A pure subgroup of a divisible group is divisible.

9.57. Give examples of a group: (a) in which the intersection of two pure subgroups is not pure; (b) in which the subgroup generated by two pure subgroups is not pure.

****9.58.** Every ascending union of pure subgroups is pure.

****9.59.** Show that an ascending union of direct summands need not be a direct summand. (HINT: Consider the group $G = \prod \sigma(p)$, and let $S_n = \sigma(p_1) \oplus \cdots \oplus \sigma(p_n)$.)

****9.60.** Let S be pure in G and let $y \in G/S$. Show that y can be lifted to $x \in G$, where x and y have the same order.

Lemma 9.28 *Let T be pure in G. If $T \subset S \subset G$ and S/T is pure in G/T, then S is pure in G.*

Proof If $x \in G$, we shall denote its coset in G/T by $\bar{x}$. Suppose that $ng = s$, where $s \in S$. Then $n\bar{g} = \bar{s}$, so that the purity of S/T guarantees an element $\bar{s}' \in S/T$ with $n\bar{s}' = \bar{s}$. Lifting this equation to G gives

$$ns' - s = t, \qquad \text{for some } t \in T.$$

Hence $ns' - ng = t$, so the purity of T yields $t' \in T$ with $ns' - ng = nt'$. Juggling, we obtain $s = n(s' - t')$; since $T \subset S$, $s' - t' \in S$, and so S is pure in G. ∎

Lemma 9.29 *A p-primary group G that is not divisible contains a pure cyclic subgroup.*

Proof Suppose there is an $x \in G[p]$ that is divisible by p^k but not by p^{k+1}; let $p^k y = x$. We let the reader prove that $\langle y \rangle$ is pure in G (Exercise 9.12 says that one need only check powers of p).

We may, therefore, assume that each $x \in G[p]$ is divisible by every power of p. We shall prove, by induction on k, that if $p^k x = 0$, then x is divisible by p. If $k = 1$, $x \in G[p]$, and our claim is certainly true. Suppose $p^{k+1}x = 0$. If $y = p^k x$, then $y \in G[p]$; hence, there is an element $z \in G$ with

$$p^{k+1}z = y = p^k x.$$

Thus $p^k(pz - x) = 0$; by induction, there is a $w \in G$ with $pw = pz - x$. Therefore

$$x = p(z - w),$$

as desired. We have shown that $G = pG$, so we know by Exercise 9.29 that G is divisible. This contradiction completes the proof. ∎

Definition A subset X of nonzero elements of a group G is **independent** in case $\sum m_\alpha x_\alpha = 0$ implies each $m_\alpha x_\alpha = 0$, where $x_\alpha \in X$ and $m_\alpha \in Z$.

Our earlier definition of independence (in a vector space) had the m_α scalars in a field, but more important, the conclusion there was that each $m_\alpha = 0$; here we pay our respects to the elements of finite order and conclude only that each $m_\alpha x_\alpha = 0$.

Lemma 9.30 *A set X of nonzero elements of G is independent if and only if*

$$\langle X \rangle = \sum_{x \in X} \langle x \rangle.$$

Proof Let $x_0 \in X$ and let $y \in \langle x_0 \rangle \cap \langle X - \{x_0\} \rangle$. Then $y = mx_0$ and $y = \sum m_\alpha x_\alpha$, where each $x_\alpha \neq x_0$. Therefore,

$$-mx_0 + \sum m_\alpha x_\alpha = 0;$$

thus each term is 0, by independence. Hence, $0 = mx_0 = y$.

The proof of the converse is left to the reader. ∎

Definition A subset X of G is **pure-independent** if X is independent and $\langle X \rangle$ is a pure subgroup of G.

Lemma 9.31 *Let G be a p-primary group. If X is a maximal pure-independent subset of G (i.e., if X is contained in no larger such), then $G/\langle X \rangle$ is divisible.*

Proof By Lemma 9.29, if $G/\langle X \rangle$ is not divisible, it contains a pure cyclic subgroup $\langle \bar{y} \rangle$. Since $\langle X \rangle$ is pure in G, $\bar{y}$ may be lifted to an element $y \in G$, where y and $\bar{y}$ have the same order (Exercise 9.60). We claim that $X^* = \{X, y\}$ is pure-independent, which will contradict the maximality of X. First of all,

$$\langle X \rangle \subset \langle X^* \rangle \subset G$$

and $\langle X^* \rangle / \langle X \rangle = \langle \bar{y} \rangle$, which is pure in $G/\langle X \rangle$. By Lemma 9.28, $\langle X^* \rangle$ is pure in G. Secondly, suppose

$$my + \sum m_\alpha x_\alpha = 0, \qquad x_\alpha \in X,\ m_\alpha, m \in Z.$$

In $G/\langle X\rangle$, this equation becomes $m\bar{y} = 0$. Since y and $\bar{y}$ have the same order, $my = 0$; since X is independent, each $m_\alpha x_\alpha = 0$. Hence, X^* is independent, and thus pure-independent. ∎

Definition Let G be a torsion group. A subgroup B of G is a **basic subgroup** of G in case:

(i) B is a direct sum of cyclic groups.
(ii) B is pure in G.
(iii) G/B is divisible.

Theorem 9.32 (Kulikov) *Every torsion group G contains a basic subgroup.*

Proof If we show that every p-primary group has a basic subgroup, then it follows from the primary decomposition that every torsion group has a basic subgroup. Assume, therefore, that G is p-primary.

If G is divisible, then $B = 0$ is a basic subgroup. If G is not divisible, then G does contain pure-independent subsets (Lemma 9.29). Since both purity and independence are preserved by ascending unions, so is pure-independence. Therefore, Zorn's lemma may be applied to provide a maximal pure-independent subset X of G. The previous two lemmas show that $B = \langle X\rangle$ is a basic subgroup. ∎

Corollary 9.33 *Every torsion group is an extension of a direct sum of cyclic groups by a divisible group.*

Corollary 9.34 (Prüfer) *Let G be a group of bounded order, i.e., $nG = 0$ for some integer $n > 0$. Then G is a direct sum of cyclic groups.*

Proof Let B be a basic subgroup of G. Then G/B is divisible, and $n(G/B) = 0$. It follows that $G/B = 0$, i.e., $B = G$. ∎

We have already classified divisible groups. If we can classify direct sums of cyclic groups, then we have classified all torsion groups modulo the extension problem. The question we ask is: If G and H are each direct sums of cyclic groups, when is $G \cong H$? The answer is essentially the same as that for finite groups given in Chapter 4. Recall that if H is a group with $pH = 0$, then it may be considered a vector space over Z_p with dimension $d(H)$. We have proved that when G is a finite p-primary group, then the number of cyclic summands of order p^n occurring in a decomposition of G is given by $d(p^{n-1}G/p^nG) - d(p^nG/p^{n+1}G)$. This formula does not generalize to infinite groups because one cannot subtract infinite cardinals.

Assume $G = \sum \sigma_i$ is a direct sum of p-primary cyclic groups. One can still prove that $d(p^nG/p^{n+1}G)$ is the number of cyclic summands of order $\geq p^{n+1}$. How can we distinguish between those elements in p^nG coming from σ_i of order p^{n+1} and those elements in p^nG coming from σ_i of larger order? Let a_i be a generator of σ_i. If $|\sigma_i| = p^{n+1}$, then p^na_i has order p; if $|\sigma_i| > p^{n+1}$, then p^na_i does not have order p. This elementary observation suggests that we amend our original idea by replacing p^nG by $p^nG \cap G[p]$, i.e., we shall be interested only in elements of order p.

Definition $U\{n, G\} = d(p^nG \cap G[p]/p^{n+1}G \cap G[p])$.

Of course, the cardinal number $U\{n, G\}$ may be infinite.

Theorem 9.35 *If G is a direct sum of p-primary cyclic groups, then the number of cyclic summands of order p^{n+1} is $U\{n, G\}$.*

Proof Let B_k denote the direct sum of all those cyclic summands of order p^k occurring in a given decomposition of G, and let b_k be the number of summands in B_k (b_k may be 0 or infinite). Thus,

$$G = B_1 \oplus B_2 \oplus \cdots \oplus B_k \oplus \cdots.$$

Now it is easy to see that

$$G[p] = B_1 \oplus pB_2 \oplus \cdots \oplus p^{k-1}B_k \oplus \cdots$$

and

$$p^nG = p^nB_{n+1} \oplus \cdots \oplus p^nB_k \oplus \cdots.$$

Therefore, for all n,

$$p^nG \cap G[p] = p^nB_{n+1} \oplus p^{n+1}B_{n+2} \oplus \cdots,$$

and so

$$p^nG \cap G[p]/p^{n+1}G \cap G[p] \cong p^nB_{n+1}.$$

Hence $U\{n, G\} = b_{n+1}$, as desired. ∎

EXERCISES

9.61. Let G and H be direct sums of p-primary cyclic groups. Prove that $G \cong H$ if and only if $U\{n, G\} = U\{n, H\}$ for all $n \geq 0$.

9.62. Prove that $t(\prod_{n=1}^{\infty} \sigma(p^n))$ is not a direct sum of cyclic groups.

9.63. If G is a p-primary group, then any two basic subgroups of G are isomorphic.

9.64. Give an example of a p-primary group G containing basic subgroups B_1 and B_2 for which $G/B_1 \ncong G/B_2$.

There is one instance when one can guarantee that a pure subgroup is a direct summand.

Theorem 9.36 *Let S be a pure subgroup of G with $nS = 0$ for some $n > 0$; then S is a direct summand of G.*

Proof Let $\pi: G \to G/(S + nG)$ be the natural map. Clearly, this quotient is of bounded order (n times it is 0), so that it is a direct sum of cyclic groups, by Prüfer's theorem. Let $G/(S + nG) = \sum \sigma(r_\alpha)$, where $\bar{x}_\alpha$ is a generator of $\sigma(r_\alpha)$. For each α, lift $\bar{x}_\alpha$ to $x_\alpha \in G$. Then $r_\alpha x_\alpha \in S + nG$, and so

$$r_\alpha x_\alpha = s_\alpha + nh_\alpha,$$

where $s_\alpha \in S$ and $h_\alpha \in G$. Now r_α divides n, so

$$s_\alpha = r_\alpha\left(x_\alpha - \frac{n}{r_\alpha} h_\alpha\right).$$

Since S is pure, there is an element $s'_\alpha \in S$ with $s_\alpha = r_\alpha s'_\alpha$. Set

$$y_\alpha = x_\alpha - s'_\alpha.$$

We have lifted and adjusted so that $r_\alpha y_\alpha = nh_\alpha$ and $\pi(y_\alpha) = \bar{x}_\alpha$. Let $K = \langle nG, \text{the } y_\alpha \rangle$; we claim that $G = S \oplus K$.

(i) $S \cap K = \{0\}$. Let $x \in S \cap K$. Since $x \in K$, $x = \sum m_\alpha y_\alpha + nh$; since $x \in S$, $\pi(x) = 0$. Hence, $0 = \sum m_\alpha \bar{x}_\alpha$, so that r_α divides m_α for each α. But we know that $r_\alpha y_\alpha \in nG$, so that surely $m_\alpha y_\alpha \in nG$. Therefore, $x = \sum m_\alpha y_\alpha + nh \in nG$; since S is pure, there is an element $s' \in S$ with $x = ns'$. But $nS = 0$; consequently $0 = ns' = x$.

(ii) $S + K = G$. If $g \in G$, then $\pi(g) = \sum m_\alpha \bar{x}_\alpha$. Since $\pi(\sum m_\alpha y_\alpha) = \sum m_\alpha \bar{x}_\alpha$, $g - \sum m_\alpha y_\alpha = s + nh \in S + nG$. Therefore, $g = s + (nh + \sum m_\alpha y_\alpha) \in S + K$. ∎

Corollary 9.37 *If tG is of bounded order, then tG is a direct summand of G. In particular, tG is a direct summand if it is finite.*

Corollary 9.38 *An indecomposable abelian group G is either torsion or torsion-free.*

Proof Let us suppose that G is an indecomposable group that is neither torsion nor torsion-free, i.e., tG is a proper subgroup of G.

Now tG is not divisible, lest it be a summand, and so tG contains a pure cyclic subgroup σ, by Lemma 9.28. It follows from Theorem 9.36 that σ is a summand of G, a contradiction. ∎

EXERCISES

9.65. Let G be a group that is not torsion-free. If tG is not divisible, then G has a finite cyclic direct summand.

9.66. A torsion group is indecomposable if and only if it is isomorphic to a subgroup of $\sigma(p^\infty)$ for some prime p.

9.67. Show that $\prod_p \sigma(p)$ is not a direct sum of (possibly infinitely many) indecomposable groups.

***9.68.** **(Kaplansky)** In the following, G is an infinite abelian group.

(a) If every proper subgroup of G is finite, then $G \cong \sigma(p^\infty)$ for some p;
(b) If G is isomorphic to every proper subgroup, then $G \cong Z$;
(c) If G is isomorphic to every proper quotient, then $G \cong \sigma(p^\infty)$ for some p;
(d) If every proper quotient of G is finite, then $G \cong Z$.

A complete classification of all countable torsion groups exists; it is due to Ulm and Zippin. The proof presupposes a knowledge of ordinal and cardinal numbers; the interested reader is referred to the books of Fuchs, Griffith, and Kaplansky (see Bibliography) for details.

TORSION-FREE GROUPS

In this section we classify a restricted class of torsion-free groups: those of rank 1. At present, there is not even an adequate classification of the groups of the finite rank, although it is known that counterexamples to almost every conjecture about direct sums can be found among these groups. We refer the reader to Fuchs, *Infinite Abelian Groups*, for further information.

When we restrict our attention to torsion-free groups, our two previous definitions of independence are equivalent, for $mx = 0$ if and only if $m = 0$ or $x = 0$.

Definition The **rank** of a torsion-free group G is the number of elements in a maximal independent subset of G.

Observe that a free abelian group is torsion-free, and our two notions of rank coincide for these groups. Exercise 9.70 below shows that the rank of a group is well defined.

EXERCISES

***9.69.** Every torsion-free group G can be imbedded in a vector space V over $\mathbf{Q}$. (HINT: First imbed G in a divisible group D and then consider the natural map $D \to D/tD$.)

***9.70.** A torsion-free group G has rank at most r if and only if G can be imbedded in an r-dimensional vector space over $\mathbf{Q}$. Conclude that rank is well defined in the sense that any two maximal independent subsets of G have the same number of elements.

9.71. Let $0 \to A \to B \to C \to 0$ be an exact sequence of torsion-free groups. Then

$$\text{rank } A + \text{rank } C = \text{rank } B.$$

Conclude that any torsion-free group of rank 1 is indecomposable.

Our starting point in examining the groups of rank 1 is that each of them is isomorphic to a subgroup of $\mathbf{Q}$ (Exercise 9.70). Let us first present three nonisomorphic subgroups of $\mathbf{Q}$.

G_1: All rationals whose denominator is square-free.
G_2: All *dyadic rationals*, i.e., all rationals of the form $m/2^k$.
G_3: All rationals whose decimal expansion is finite.

It is instructive to prove that no two of the groups G_1, G_2, G_3, Z, $\mathbf{Q}$ are isomorphic.

A perceptive observation is that each of these groups can be described by the numbers that are allowed to be denominators. (G_3 may be alternatively described as those rationals whose denominators are restricted to be powers of 10.)

Let $p_1, p_2, \cdots, p_n, \cdots$ be the sequence of primes.

Definition A **characteristic** is a sequence

$$(k_1, k_2, \cdots, k_n, \cdots)$$

where each k_n is a nonnegative integer or the symbol ∞.

If G is a subgroup of $\mathbf{Q}$ and if $x \in G$ is nonzero, then x determines a characteristic in the following way: What is the highest power of p_n that

divides x in G? That is, for which nonnegative integers k is there an element $y \in G$ satisfying

$$p_n^k y = x?$$

If there is a largest such exponent k, we call it k_n; if there is no largest such exponent, we set $k_n = \infty$.[5]

It is convenient to write each nonzero integer as a formal infinite product $\prod p_i^{\alpha_i}$, where the p_i range over all the primes and $\alpha_i \geq 0$ (of course almost all the $\alpha_i = 0$). Let $m = \prod p_i^{\alpha_i}$ and $n = \prod p_i^{\beta_i}$ be given integers. If $a \in G$ has characteristic $(k_1, k_2, \cdots)$, then the definition of characteristic says that there is an $x \in G$ satisfying $mx = na$ if and only if $\alpha_i \leq k_i + \beta_i$ for all i (by convention, $\infty + \beta_i = \infty$).

Each of the five groups of rank 1 exhibited above contains $x = 1$. Its characteristic in each group is

$$\begin{aligned} Z&: \quad (0, 0, 0, \cdots);\\ \mathbf{Q}&: \quad (\infty, \infty, \infty, \cdots);\\ G_1&: \quad (1, 1, 1, \cdots);\\ G_2&: \quad (\infty, 0, 0, \cdots);\\ G_3&: \quad (\infty, 0, \infty, 0, 0, 0, \cdots). \end{aligned}$$

Unfortunately, distinct nonzero elements of the same group G may give rise to distinct characteristics. For example, if $G = G_2 =$ dyadic rationals, the characteristic of 1 is

$$(\infty, 0, 0, \cdots)$$

while the characteristic of $126 = 2 \cdot 3^2 \cdot 7$ is

$$(\infty, 2, 0, 1, 0, 0, \cdots).$$

We are led to the following definition.

Definition Two characteristics are **equivalent** if

(i) they have ∞ in the same coordinates;
(ii) they differ in at most a finite number of coordinates.

It is easy to check that this is an equivalence relation; an equivalence class of characteristics is called a **type**.

Lemma 9.39 *Let G be a subgroup of $\mathbf{Q}$ and let x and x' be nonzero elements of G. Then the characteristics of x and of x' are equivalent.*

Proof First of all, if $x' = mx$ for some integer m, then the characteristics of x and x' are equivalent: x' is divisible by every power of p_i

[5] k_n is called the **p_n-height** of x.

that divides x (plus only a few more); x' is divisible by every power of p_i if and only if x is.

Let us now pass to the general case. Since G is a subgroup of $\mathbf{Q}$, there are integers m and n such that

$$mx = nx'.$$

The characteristic of x is equivalent to that of $mx = nx'$, which is equivalent to that of x'. ∎

As a result of this lemma, if G is a group of rank 1 (i.e., a subgroup of $\mathbf{Q}$), we may define the **type** of G, $\tau(G)$, as the type of a nonzero element of G.

Theorem 9.40 *Let G and G' be torsion-free groups of rank* 1. *Then $G \cong G'$ if and only if $\tau(G) = \tau(G')$.*

Proof Suppose $f: G \to G'$ is an isomorphism. If $x \in G$ is nonzero, then one verifies easily that x and $f(x)$ have the same characteristic. Therefore, $\tau(G) = \tau(G')$.

Assume that $\tau(G) = \tau(G')$ and (without loss of generality) that G and G' are subgroups of $\mathbf{Q}$. If a and a' are nonzero elements of G and G', respectively, then their characteristics $(k_1, k_2, \cdots)$ and $(k_1', k_2', \cdots)$ differ in only a finite number of places. If we agree that the notation $\infty - \infty$ means 0, then we may define a rational number λ by

$$\lambda = \prod p_i^{k_i - k_i'}$$

It follows from the definition of equivalence and our convention concerning ∞ that almost all the $k_i - k_i' = 0$.

Define $f: G \to \mathbf{Q}$ by $f(x) = \mu x$, where $\mu = \lambda a'/a$. Note that distributivity implies that f is a homomorphism. Now a rational x is in G if and only if there are integers $m = \prod p_i^{\alpha_i}$ and $n = \prod p_i^{\beta_i}$ with $mx = na$ and $\alpha_i \leq \beta_i + k_i$ for all i; a rational y is in G' if and only if there are integers m and n with $my = na'$ and $\alpha_i \leq \beta_i + k_i'$ for all i. We claim that image $f \subset G'$. If $x \in G$, then $mx = na$ and $\alpha_i \leq \beta_i + k_i$; hence, $m(\mu x) = n\mu a = (n\lambda)a'$. Since $\alpha_i \leq (\beta_i + k_i - k_i') + k_i'$, it follows that $\mu x = f(x) \in G'$. In a similar manner, the reader may see that if $g: G' \to \mathbf{Q}$ is defined by $g(x') = \mu^{-1}x'$, then image $g \subset G$. Therefore, f and g are inverse, and $G \cong G'$. ∎

Theorem 9.41 *If τ is a type, then there exists a group G of rank* 1 *with $\tau(G) = \tau$.*

Proof Let $(k_1, k_2, \cdots, k_n, \cdots)$ be a characteristic in τ. We define a group G as the subgroup of $\mathbf{Q}$ generated by all rationals of the form $1/m$, where, for all n, p_n^t divides m if and only if $t \leq k_n$. It is easy to check that G is a group of rank 1 and that the element 1 in G has the given characteristic. ∎

EXERCISES

9.72. If G is a subring of $\mathbf{Q}$, then G is also a torsion-free additive group of rank 1. Prove that the characteristic of 1 in G has only 0 and ∞ as entries.

9.73. Prove that two distinct subrings of $\mathbf{Q}$ are not isomorphic as additive groups.

9.74. There are uncountably many nonisomorphic subgroups of $\mathbf{Q}$.

9.75. If G is any abelian group, the set of all endomorphisms of G is a ring if we define multiplication as composition and addition by $(f + g)(x) = f(x) + g(x)$.

If G is torsion-free of rank 1, prove that the endomorphism ring of G is isomorphic to a subring of $\mathbf{Q}$. (HINT: Use the injective property of $\mathbf{Q}$.)

9.76. Give an example of two nonisomorphic groups having isomorphic endomorphism rings.

9.77. Let G and H be torsion-free of rank 1. If there are nonzero elements $x \in G$ and $y \in H$ with $G/\langle x \rangle \cong H/\langle y \rangle$, prove that $G \cong H$. What is the relation between $G/\langle x \rangle$ and the characteristic of x?

Every theorem in this chapter may be generalized to modules over principal ideal domains; they may be then interpreted as an analysis of linear transformations operating on infinite-dimensional spaces.

GROTHENDIECK GROUPS

In this last section, we give a construction that will allow us to state the Jordan-Hölder and Remak-Krull-Schmidt theorems in a very elegant way. The basic idea is that the isomorphism classes of a set of groups form a semigroup under the binary operation of direct product; we force this semigroup to be a group.

Definition Let $\mathscr{S}$ be a set of (not necessarily abelian) groups. The **Grothendieck group** of $\mathscr{S}$, denoted $K(\mathscr{S})$, is the abelian group with generators the elements of $\mathscr{S}$ which are subject to the relations

$$A + C = B \quad \text{if} \quad B \cong A \times C.$$

In more detail, let $\mathscr{F}$ be the free abelian group with basis $\mathscr{S}$, and let $\mathscr{R}$ be the subgroup of $\mathscr{F}$ generated by all elements of the form $A + C - (A \times C)$. Then $K(\mathscr{S}) = \mathscr{F}/\mathscr{R}$. If $A \in \mathscr{S}$, then we denote $A + \mathscr{R}$ by $[A]$. Note that $A \cong A'$ implies $[A] = [A']$ if $\{1\} \in \mathscr{S}$.

Every element in $K(\mathscr{S})$ has an inverse: If $A \in \mathscr{S}$, what is $-[A]$? That element which, when added to $[A]$, gives 0! The point is that not every element in $K(\mathscr{S})$ is of the form $[B]$ for some $B \in \mathscr{S}$. Indeed, this is why we give the construction above.

In what follows, we assume that $\{1\} \in \mathscr{S}$ and that $\mathscr{S}$ is closed under finite direct products.

Lemma 9.42 *Every element of $K(\mathscr{S})$ has the form $[A] - [B]$, where $A, B \in \mathscr{S}$.*

Proof If $x \in K(\mathscr{S})$, then $x = \sum_{i=1}^{n} A_i - \sum_{j=1}^{m} B_j + \mathscr{R}$, where $A_i, B_j \in \mathscr{S}$ (we do not need integer coefficients, for we are willing to repeat an A_i or a B_j several times if necessary). In bracket notation,

$$\begin{aligned} x &= \sum [A_i] - \sum [B_j] \\ &= [A_1 \times \cdots \times A_n] - [B_1 \times \cdots \times B_m]. \end{aligned}$$ ∎

Lemma 9.43 *Let $A, B \in \mathscr{S}$. Then $[A] = [B] \in K(\mathscr{S})$ if and only if there is a group $C \in \mathscr{S}$ with $A \times C \cong B \times C$.*

Proof If $A \times C \cong B \times C$, then

$$[A] + [C] = [A \times C] = [B \times C] = [B] + [C].$$

Since $K(\mathscr{S})$ is a group, we may cancel $[C]$ to obtain $[A] = [B]$.

Suppose, conversely, that $[A] = [B]$. Then $A - B \in \mathscr{R}$, so there is an equation in $\mathscr{F}$:

$$A - B = \sum (X_i' + X_i'' - X_i) - \sum (Y_j' + Y_j'' - Y_j),$$

where $X_i = X_i' \times X_i''$ and $Y_j = Y_j' \times Y_j''$. Now transpose to eliminate negative coefficients:

$$A + \sum X_i + \sum (Y_j' + Y_j'') = B + \sum (X_i' + X_i'') + \sum Y_j.$$

Since this equation is a relation among the members of a basis of $\mathscr{F}$, each term appearing on one side occurs on the other side and with

equal frequency. It follows that the direct product of all the groups on the left is isomorphic to the direct product of those on the right. If $X = \prod X_i$, $X' = \prod X'_i$, etc., then

$$A \times X \times Y' \times Y'' \cong B \times X' \times X'' \times Y.$$

But, by definition, $X \cong X' \times X''$ and $Y \cong Y' \times Y''$, so if we define $C = X \times Y' \times Y''$, we have $A \times C \cong B \times C$. ∎

Theorem 9.44 *Let $\mathcal{S}$ be a set of groups whose members satisfy either chain condition. If $\mathcal{S}$ is closed under finite direct products and direct factors, then $K(\mathcal{S})$ is free abelian with basis $X = \{[A]: A \in \mathcal{S}$ and A indecomposable$\}$ if and only if the Remak-Krull-Schmidt theorem holds for each member of $\mathcal{S}$.*

Proof Assume that $K(\mathcal{S})$ is free abelian with basis X. This means, among other things, that if A and B are indecomposable and $A \ncong B$, then $[A] \neq [B]$ and $[A] \neq 0$. If $G \in \mathcal{S}$, Lemma 4.30 says G is a direct product of finitely many indecomposables. Suppose

$$G \cong A_1 \times \cdots \times A_n \cong B_1 \times \cdots \times B_m,$$

where the A_i and B_j are indecomposable. In $K(\mathcal{S})$, we have

$$\sum_{i=1}^{n} [A_i] = [A_1 \times \cdots \times A_n] = [G] = [B_1 \times \cdots \times B_m] = \sum_{j=1}^{m} [B_j].$$

By independence of X, $n = m$ and there is a permutation σ of $\{1, \cdots, n\}$ with $[A_i] = [B_{\sigma(i)}]$ for all i. By our initial remarks, $A_i \cong B_{\sigma(i)}$ for all i, as desired.

To prove the converse, let us first show that if A and B are non-isomorphic indecomposable groups, then $[A] \neq [B]$ and $[A] \neq 0$. If $[A] = [B]$, then Lemma 9.43 provides a group $C \in \mathcal{S}$ with $A \times C \cong B \times C$. By Exercise 4.70 (essentially Remak-Krull-Schmidt), we may cancel C to obtain $A \cong B$. We reach a similar contradiction if we assume $[A] = 0$.

If $G \in \mathcal{S}$, then $G = A_1 \times \cdots \times A_n$, where each A_i is indecomposable; in $K(\mathcal{S})$, this becomes $[G] = [A_1] + \cdots + [A_n]$. Therefore, the set of $[A]$ with A indecomposable generates $K(\mathcal{S})$.

To show independence, suppose

$$\sum_{i=1}^{n} [A_i] - \sum_{j=1}^{m} [B_j] = 0,$$

where A_i and B_j are indecomposable groups. Then

$$[A_1 \times \cdots \times A_n] = \sum [A_i] = \sum [B_j] = [B_1 \times \cdots \times B_m].$$

By Lemma 9.43, there is a group $C \in \mathscr{S}$ with

$$A_1 \times \cdots \times A_n \times C \cong B_1 \times \cdots \times B_m \times C;$$

we cancel C, by Exercise 4.70, to obtain

$$A_1 \times \cdots \times A_n \cong B_1 \times \cdots \times B_m.$$

The Remak-Krull-Schmidt theorem gives $n = m$ and a permutation σ of $\{1, \cdots, n\}$ so that $A_i \cong B_{\sigma(i)}$ for all i. The original dependence relation is thus the relation with all coefficients 0 (after we collect terms). ∎

EXERCISES

9.78. If $\mathscr{S}$ is a family of groups closed under infinite direct sums (or infinite direct products), then $K(\mathscr{S}) = 0$.

9.79. Compute $K(\mathscr{S})$ when $\mathscr{S}$ is the family of all finitely generated abelian groups.

****9.80.** If $1 \to X_i' \to X_i \to X_i'' \to 1$ is exact for $i = 1, \cdots, n$, then there is an exact sequence $1 \to \prod X_i' \to \prod X_i \to \prod X_i'' \to 1$.

****9.81.** If $1 \to L \to M \to N \to 1$ is exact, then there is an exact sequence $1 \to L \to B \times M \to B \times N \to 1$.

We now modify the definition of $K(\mathscr{S})$ by imposing different relations on the generators.

Definition Let $\mathscr{S}$ be a set of groups. The **Grothendieck group** $K^*(\mathscr{S})$ is the abelian group with generators the elements of $\mathscr{S}$, which are subject to the relations $A + C = B$ if there is an exact sequence $1 \to A \to B \to C \to 1$, i.e., if B is an extension of A by C.

We maintain the bracket notation: if $A \in \mathscr{S}$, then $[A] = A + \mathscr{R}^* \in K^*(\mathscr{S})$, where $\mathscr{R}^*$ is the subgroup of $\mathscr{F}$ generated by the relations in the definition of K^*.

Addition in $K^*(\mathscr{S})$ is interesting.

Lemma 9.45 *Let* $G = G_0 \supset G_1 \supset \cdots \supset G_n = \{1\}$ *be a normal series with factor groups* $Q_i = G_{i-1}/G_i$. *In* $K^*(\mathscr{S})$,

$$[G] = [Q_1] + [Q_2] + \cdots + [Q_n].$$

Proof The equation $Q_i = G_{i-1}/G_i$ becomes $[Q_i] = [G_{i-1}] - [G_i]$ in $K^*(\mathscr{S})$. Therefore

$$\sum_{i=1}^{n} [Q_i] = \sum_{i=1}^{n} ([G_{i-1}] - [G_i]) = [G_0] - [G_n] = [G]. \quad \blacksquare$$

It follows from this lemma that $[A_1 \times \cdots \times A_n] = [A_1] + \cdots + [A_n]$ in $K^*(\mathscr{S})$ since there is a normal series

$$A_1 \times \cdots \times A_n \supset A_2 \times \cdots \times A_n \supset \cdots \supset A_n \supset 1.$$

Lemma 9.46 *Let* $A, B \in \mathscr{S}$. *Then* $[A] = [B] \in K^*(\mathscr{S})$ *if and only if there is a group* $C \in \mathscr{S}$ *and normal series of* $A \times C$ *and* $B \times C$ *with the same factor groups.*

Proof Assume there is a group C so that there are normal series of $A \times C$ and $B \times C$ having factor groups $Q_1, \cdots, Q_n$. By Lemma 9.45, in $K^*(\mathscr{S})$,

$$[A \times C] = [Q_1] + \cdots + [Q_n] = [B \times C].$$

But $A \times C$ is an extension of A by C; hence

$$[A \times C] = [A] + [C]$$

and

$$[B \times C] = [B] + [C].$$

Therefore $[A] = [B]$.

Assuming $[A] = [B]$, we shall first show there exist groups C, U, V and exact sequences

$$1 \to U \to A \times C \to V \to 1$$

and

$$1 \to U \to B \times C \to V \to 1.$$

As in the proof of Lemma 9.43, $A - B \in \mathscr{R}^*$ gives an equation in $\mathscr{F}$:

$$A + \sum X_i + \sum (Y_j' + Y_j'') = B + \sum (X_i' + X_i'') + \sum Y_j,$$

where now we know that there are exact sequences $1 \to X_i' \to X_i \to X_i'' \to 1$ and $1 \to Y_j' \to Y_j \to Y_j'' \to 1$. Setting $X = \coprod X_i$, $X' = \coprod X_i'$, etc., we obtain as in Lemma 9.43

$$A \times X \times Y' \times Y'' \cong B \times X' \times X'' \times Y.$$

Let C be a group isomorphic to both of these. Observe that there is an exact sequence $1 \to X' \to X \to X'' \to 1$ arising from the definition of X_i', X_i, X_i''. This gives rise to exact sequences

$$1 \to X' \times Y' \to X \times Y' \to X'' \to 1$$

and

$$1 \to X' \times Y' \to A \times X \times Y' \times Y'' \to A \times X'' \times Y'' \to 1.$$

Since the middle term is C, this gives rise to a third exact sequence (Exercise 9.81)

$$1 \to X' \times Y' \to B \times C \to B \times A \times X'' \times Y'' \to 1.$$

If we define $U = X' \times Y'$ and $V = B \times A \times X'' \times Y''$, then we have an exact sequence $1 \to U \to B \times C \to V \to 1$. We let the reader perform similar manipulations on $B \times X' \times X'' \times Y$ to get an exact sequence $1 \to U \to A \times C \to V \to 1$.

To complete the proof, note that we have normal series

$$A \times C \supset U \supset \{1\} \qquad \text{and} \qquad B \times C \supset U \supset \{1\},$$

each of which has factor groups U and V. ∎

Theorem 9.47 *Let $\mathcal{S}$ be the set of all finite groups. Then $K^*(\mathcal{S})$ is free abelian with basis $X = \{[A]: A \in \mathcal{S}$ and A is simple$\}$ if and only if the Jordan-Hölder theorem holds for each member of $\mathcal{S}$.*

Proof Assume that $K^*(\mathcal{S})$ is free abelian with basis X. This means, among other things, that if A and B are simple and $A \ncong B$, then $[A] \neq [B]$ and $[A] \neq 0$. If $G \in \mathcal{S}$, then G has a composition series. Suppose G has two composition series with factor groups $A_1, \cdots, A_n$ and $B_1, \cdots, B_m$. By Lemma 9.45, we have in $K^*(\mathcal{S})$,

$$[A_1] + \cdots + [A_n] = [G] = [B_1] + \cdots + [B_m];$$

moreover, the A_i and B_j are simple, being factor groups of composition series. By independence of X, $n = m$ and there is a permutation σ of $\{1, \cdots, n\}$ with $[A_i] = [B_{\sigma(i)}]$ for all i. By our initial remarks, $A_i \cong B_{\sigma(i)}$ for all i, as desired.

To prove the converse, we first show that if A and B are nonisomorphic simple groups, then $[A] \neq [B]$ and $[A] \neq 0$. If $[A] = [B]$, then there is a finite group C and normal series

$$A \times C \supset G_1 \supset \cdots \supset G_n = 1$$

and

$$B \times C \supset H_1 \supset \cdots \supset H_m = 1$$

having the same factor groups. By Exercise 6.16, $A \times C$ and $B \times C$ have the same composition factors. But, since A is simple, the composition factors of $A \times C$ are those of C together with A; similarly, we know the composition factors of $B \times C$. By the Jordan-Hölder theorem, we must have $A \cong B$. A similar argument shows that $[A] \neq 0$ when A is simple.

If $G \in \mathscr{S}$ has composition factors $A_1, \cdots, A_n$, then each A_i is simple and, by Lemma 9.45, $[G] = [A_1] + \cdots + [A_n]$. Therefore the "simple" $[A]$'s generate $K^*(\mathscr{S})$.

To show independence, suppose

$$\sum_{i=1}^{n} [A_i] - \sum_{j=1}^{m} [B_j] = 0,$$

where A_i and B_j are simple groups. Then

$$[A_1 \times \cdots \times A_n] = [B_1 \times \cdots \times B_m]$$

(by the remark after Lemma 9.45); there thus exists a finite group C such that $A_1 \times \cdots \times A_n \times C$ and $B_1 \times \cdots \times B_m \times C$ have normal series with the same factor groups. By Exercise 6.16, $A_1 \times \cdots \times A_n \times C$ and $B_1 \times \cdots \times B_m \times C$ have the same composition factors. But the composition factors of $A_1 \times \cdots \times A_n \times C$ are $A_1, \cdots, A_n$, together with the composition factors of C. The Jordan-Hölder theorem gives $n = m$ and a permutation σ of $\{1, \cdots, n\}$ with $A_i \cong B_{\sigma(i)}$ for all i. The original dependence relation is thus the relation with all coefficients 0 (after we collect terms). ∎

The idea that an extension of A by C is a product has now been formalized so that it is, in fact, a product (rather, a sum) of two elements in a group.

EXERCISES

9.82. Prove that the function $K(\mathscr{S}) \to K^*(\mathscr{S})$ that sends $G + \mathscr{R}$ into $G + R^*$ is a homomorphism onto.

9.83. Compute $K^*(\mathscr{S})$ when $\mathscr{S}$ is the set of all finite abelian groups.

9.84. If $\mathscr{S}$ is the set of all finitely generated abelian groups, prove that $K^*(\mathscr{S}) \cong Z$.

CHAPTER

10

Homological Algebra

THE HOM FUNCTORS

In Chapter 1 we raised the twin questions of describing groups and describing homomorphisms, but our emphasis to this point has been upon groups. In this chapter we focus upon homomorphisms with the ultimate goal of computing some groups of extensions. We restrict ourselves here to abelian groups, but the reader should regard this study as an introduction to the more general theory that includes $H_\theta^2(Q, K)$ and the Schur-Zassenhaus lemma (Theorem 7.19).

For the remainder of this chapter, *group* means abelian group.

The fundamental abstraction we need is that of a functor, which we define at once. Let $\mathscr{A}$ be the class of all abelian groups; if $A \in \mathscr{A}$, we write 1_A for the identity map on A.

Definition A function $T: \mathscr{A} \to \mathscr{A}$ is a **covariant** (additive) **functor** if whenever $\alpha: A \to B$ is a homomorphism, there is a homomorphism $T(\alpha): T(A) \to T(B)$ such that:

(i) $T(1_A) = 1_{T(A)}$;
(ii) $T(\beta\alpha) = T(\beta)T(\alpha)$ (where $\beta: B \to C$);
(iii) $T(\alpha + \gamma) = T(\alpha) + T(\gamma)$ (where $\gamma: A \to B$).

The following construction gives an important example of a covariant functor.

Definition Let G be a fixed group. Define a function $T: \mathscr{A} \to \mathscr{A}$ by $T(A) = \mathrm{Hom}(G, A)$, the group of all homomorphisms of G into A under the binary operation

$$(f + g)(x) = f(x) + g(x).$$

If $\alpha: A \to B$, define $T(\alpha): T(A) \to T(B)$ as follows: If $f \in T(A) = \mathrm{Hom}(G, A)$, then $f: G \to A$; define $T(\alpha)(f) = \alpha f: G \to B$.

Theorem 10.1 *If G is a group, then $T = \mathrm{Hom}(G, \ )$ is a covariant functor.*

Proof The verifications of the axioms are automatic. ∎

Definition A function $T: \mathscr{A} \to \mathscr{A}$ is a **contravariant** (additive) **functor** if whenever $\alpha: A \to B$ is a homomorphism, there is a homomorphism $T(\alpha): T(B) \to T(A)$ such that:

(i) $T(1_A) = 1_{T(A)}$;
(ii) $T(\beta\alpha) = T(\alpha)T(\beta)$ (where $\beta: B \to C$);
(iii) $T(\alpha + \gamma) = T(\alpha) + T(\gamma)$ (where $\gamma: A \to B$).

Observe that the main difference between covariant and contravariant functors is that the latter changes the direction of arrows.

Theorem 10.2 *If K is a group, then $S = \mathrm{Hom}(\ , K)$ is a contravariant functor.*

Proof If $A \in \mathscr{A}$, define $S(A) = \mathrm{Hom}(A, K)$. If $\alpha: A \to B$, define $S(\alpha): \mathrm{Hom}(B, K) \to \mathrm{Hom}(A, K)$ by $S(\alpha)(g) = g\alpha$, where $g: B \to K$. The remaining details are left to the reader. ∎

We now have examples of functors of either variance; if we use the term "functor" without a modifier, then what we say is to be true for every functor, contra or co.

EXERCISES

10.1. The identity functor $J: \mathscr{A} \to \mathscr{A}$, defined by $J(A) = A$ and $J(\alpha) = \alpha$, is a covariant functor.

10.2. The **zero map** $\alpha: A \to B$ is the homomorphism defined by $\alpha(x) = 0$ for all $x \in A$. If T is a functor, then $T(\alpha) = 0$ for every zero map α, and $T(0) = 0$, where 0 is the zero group.

10.3. If $\alpha\colon A \to B$ is an isomorphism and T is a functor, then $T(\alpha)$ is an isomorphism.

10.4. Which of the following are functors?

(a) $T(G) = tG$ and $T(f) = f|tG$.
(b) $T(G) = dG$ and $T(f) = f|dG$.
(c) $T(G) = G \oplus A$ (where A is fixed) and $T(f) = f \oplus 1_A$.
(d) $T(G) = G$ and $T(f) = -f$.

10.5. Let $A_1, A_2, \cdots, A_n$ be a finite set of groups and let T be a functor; prove that $T(\sum A_k) \cong \sum T(A_k)$. (HINT: Use Exercise 4.9.) (We shall soon see that this is false if we delete the adjective "finite".)

Definition A **short exact sequence** is an exact sequence of the form

$$0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0; \tag{*}$$

(*) is **split** in case there is a map $\delta\colon C \to B$ with $\beta\delta = 1_C$.

EXERCISES

10.6. The short exact sequence (*) is split if and only if there is a map $\gamma\colon B \to A$ with $\gamma\alpha = 1_A$.

10.7. If the short exact sequence (*) is split, then $B \cong A \oplus C$, and B is a split extension.

10.8. If T is a covariant functor and (*) is split, then $0 \to T(A) \to T(B) \to T(C) \to 0$ is a split short exact sequence (similarly for contravariant functors).

Definition A covariant functor T is **left exact** in case exactness of

$$0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C$$

implies exactness of

$$0 \to T(A) \to T(B) \to T(C);$$

a contravariant functor T is **left exact** in case exactness of

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$$

implies exactness of

$$0 \to T(C) \to T(B) \to T(A).$$

Theorem 10.3 *$S = \mathrm{Hom}(\ , K)$ is a left exact functor.*

Proof Let $A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ be exact. Since S is contravariant, we get the sequence

$$0 \longrightarrow S(C) \xrightarrow{S(\beta)} S(B) \xrightarrow{S(\alpha)} S(A),$$

i.e.,

$$0 \longrightarrow \mathrm{Hom}(C, K) \xrightarrow{S(\beta)} \mathrm{Hom}(B, K) \xrightarrow{S(\alpha)} \mathrm{Hom}(A, K),$$

which we claim is exact. It must be shown that $S(\beta)$ is one-to-one and that image $S(\beta) =$ kernel $S(\alpha)$.

(i) Kernel $S(\beta) = 0$. Suppose $f: C \to K$ and $S(\beta)f = 0$, i.e., $f\beta = 0$. Then f annihilates image $\beta = C$, since β is onto, and so $f = 0$.

(ii) Image $S(\beta) \subset$ kernel $S(\alpha)$. If $f: C \to K$, then $S(\alpha)S(\beta)(f) = S(\alpha)(f\beta) = f(\beta\alpha) = 0$, since $\beta\alpha = 0$.

(iii) Kernel $S(\alpha) \subset$ image $S(\beta)$. Suppose $S(\alpha)(g) = 0$ where $g: B \to K$; thus $g\alpha = 0$. Define $g_{\#}: C \to K$ by $g_{\#}(c) = g(b)$ where $\beta(b) = c$. Now $g_{\#}$ is well defined, for if $\beta(b') = c$, then $b - b' \in$ kernel $\beta =$ image α, so that $b - b' = \alpha(a)$ for some $a \in A$. Therefore $g(b - b') = g\alpha(a) = 0$, so that $g(b) = g(b')$. But $S(\beta)(g_{\#}) = g_{\#}\beta = g$, for if $b \in B$ and $c = \beta(b)$, then $g_{\#}\beta(b) = g_{\#}(c) = g(b)$. ∎

Theorem 10.4 *$T = \mathrm{Hom}(G, \ )$ is a left exact functor.*

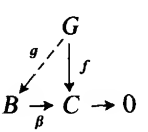
Proof Similar to the proof given above. ∎

The answer to the question of when "$\to 0$" can be tagged on the end of the functored sequence is given below.

Theorem 10.5 *A group G is free if and only if, whenever $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is exact, the sequence*

$$0 \to \mathrm{Hom}(G, A) \xrightarrow{T(\alpha)} \mathrm{Hom}(G, B) \xrightarrow{T(\beta)} \mathrm{Hom}(G, C) \to 0$$

is exact.



Proof Note that the critical assumption is that $T(\beta)$ is onto. Suppose G is free; consider the diagram above, where $f \in \mathrm{Hom}(G, C)$. By Theorem 9.20, G has the projective property; there is thus a map $g\colon G \to B$ with $\beta g = f$. But $\beta g = T(\beta)(g)$, so $f \in$ image $T(\beta)$, as desired.

If $T(\beta)$ is onto, then every $f\colon G \to C$ is of the form $T(\beta)(g)$ for some $g \in \mathrm{Hom}(G, B)$, i.e., $f = g\beta$. Since we are assuming $T(\beta)$ is onto for every exact sequence $A \to B \to C \to 0$, G has the projective property. Therefore, G is free, by Exercise 9.43. ∎

Theorem 10.6 *A group K is divisible if and only if, whenever $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is exact, the sequence*

$$0 \to \mathrm{Hom}(C, K) \to \mathrm{Hom}(B, K) \to \mathrm{Hom}(A, K) \to 0$$

is exact.

Proof Use the injective property. ∎

Theorem 10.7 *Let $\{A_j : j \in J\}$ be a family of groups; for any group K,*

$$\mathrm{Hom}(\textstyle\sum A_j, K) \cong \prod \mathrm{Hom}(A_j, K).$$

Proof For each j, let $i_j\colon A_j \to \sum A_j$ be the inclusion. Define a map $\theta\colon \mathrm{Hom}(\sum A_j, K) \to \prod \mathrm{Hom}(A_j, K)$ by $f \to (fi_j)$. Define a map ψ in the reverse direction by $(f_j) \to f$, where f is the unique map of $\sum A_j$ into K (Exercise 9.9) such that $fi_j = f_j$ for all j. A mechanical check shows that θ and ψ are inverse. ∎

We now have an example of a functor that does not preserve infinite direct sums.

Theorem 10.8 *Let $\{A_j : j \in J\}$ be a family of groups; for any group G,*

$$\mathrm{Hom}(G, \textstyle\prod A_j) \cong \prod \mathrm{Hom}(G, A_j).$$

Proof For each j, let $p_j\colon \prod A_j \to A_j$ be the jth projection. An argument dual to the argument given above, using Exercise 9.10, shows that $\theta\colon \mathrm{Hom}(G, \prod A_j) \to \prod \mathrm{Hom}(G, A_j)$, defined by $f \to (p_j f)$, is an isomorphism. ∎

Theorem 10.9 *Let $\mu\colon A \to A$ be multiplication by m, i.e., $\mu(a) = ma$ for all $a \in A$. If T is a functor, then $T(\mu)\colon T(A) \to T(A)$ is also multiplication by m.*

Proof If we let $\alpha = 1_A$, then $\mu = \alpha + \cdots + \alpha$ (m times). Since T is a functor, $T(\mu) = T(\alpha) + \cdots + T(\alpha)$ (m times) and $T(\alpha) = T(1_A) = 1_{T(A)}$. Therefore, $T(\mu)$ is also multiplication by m. ∎

We now present some sample computations to illustrate how these theorems may be used.

Example 1 For any group G, $\mathrm{Hom}(G, \mathbf{Q})$ is a vector space over $\mathbf{Q}$.

A group H is torsion-free if and only if every nonzero multiplication $\mu\colon H \to H$ is one-to-one; a group H is divisible if and only if every nonzero multiplication $\mu\colon H \to H$ is onto. Thus, a group is torsion-free and divisible, i.e., is a vector space over $\mathbf{Q}$, if and only if every nonzero multiplication is an automorphism.

If $\mu\colon \mathbf{Q} \to \mathbf{Q}$ is multiplication by $m \neq 0$, then μ is an isomorphism. If T is the functor $\mathrm{Hom}(G, \ )$, then $T(\mu)$ is an isomorphism (Exercise 10.3) that is also multiplication by m. Therefore, $T(G) = \mathrm{Hom}(G, \mathbf{Q})$ is torsion-free and divisible and hence is a vector space over $\mathbf{Q}$.

Example 2 For any group G, $\mathrm{Hom}(Z, G) \cong G$.

The map $\theta\colon \mathrm{Hom}(Z, G) \to G$, defined by $f \to f(1)$, is an isomorphism.

Example 3 For any group G, $\mathrm{Hom}(\sigma(n), G) \cong G[n]$.

Consider the exact sequence

$$0 \to Z \xrightarrow{\mu} Z \to \sigma(n) \to 0$$

where μ is multiplication by n. If we apply the functor $S = \mathrm{Hom}(\ , G)$ to this sequence, then we obtain the exact sequence

$$0 \to \mathrm{Hom}(\sigma(n), G) \to \mathrm{Hom}(Z, G) \xrightarrow{S(\mu)} \mathrm{Hom}(Z, G);$$

thus $\mathrm{Hom}(\sigma(n), G) = \text{kernel } S(\mu)$. Since $\mathrm{Hom}(Z, G) \cong G$ and $S(\mu)$ is multiplication by n, $\mathrm{Hom}(\sigma(n), G) \cong G[n]$.

Example 4 For any group G, let $G^* = \mathrm{Hom}(G, \mathbf{T})$, where $\mathbf{T}$ is the circle group; G^* is called the **character group** of G. If

$$0 \to A \to B \to C \to 0$$

is exact, then so is

$$0 \to C^* \to B^* \to A^* \to 0.$$

We are applying the functor $\mathrm{Hom}(\ , \mathbf{T})$, and $\mathbf{T}$ is a divisible group. Exactness of the sequence of character groups thus follows from Theorem 10.6.

EXERCISES

10.9. Show that $\mathrm{Hom}(A, B) = 0$ in each of the following cases:
(a) A is torsion, B is torsion-free.
(b) A is divisible, B is reduced.
(c) A is p-primary, B is q-primary, $p \neq q$.

10.10. Prove that $\mathrm{Hom}(G, \mathbf{T}) = 0$ if and only if $G = 0$.

10.11. If G is finite, prove that $G \cong G^*$, where G^* is the character group of G.

10.12. If S is a subgroup of the finite group G, set

$$S^0 = \{f \in G^* : f(s) = 0 \quad \text{for all } s \in S\}.$$

Prove that S^0 is a subgroup of G^* and that $S^0 \cong (G/S)^*$.

10.13. Let G be finite of order n and let k divide n. Prove that G has the same number of subgroups of order k as of subgroups of index k. (HINT: Prove that $S \to S^0$ is a one-to-one correspondence.)

10.14. Use character groups to give an alternative proof of Exercise 4.29.

10.15. We now present counterexamples showing that Theorems 10.7 and 10.8 cannot be extended. Let p_i denote the ith prime.

(a) If $A_i = \sigma(p_i)$, prove that $\mathrm{Hom}(\prod A_i, \mathbf{Q}) \neq 0$. Conclude that $\mathrm{Hom}(\prod A_i, \mathbf{Q})$ is isomorphic to neither $\prod \mathrm{Hom}(A_i, \mathbf{Q})$ nor $\sum \mathrm{Hom}(A_i, \mathbf{Q})$.
(b) If $A_i = \sigma(p_i)$ and $G = \sum A_i$, then $\mathrm{Hom}(G, \sum A_i)$ and $\sum \mathrm{Hom}(G, A_i)$ are not isomorphic.
(c) If $G = Z$ and $A_i = \sigma(p_i)$, show that $\mathrm{Hom}(G, \sum A_i)$ and $\prod \mathrm{Hom}(G, A_i)$ are not isomorphic.

DEFINITION OF EXT

We wish to consider all abelian extensions B of a group A by a group C, and so we recall the following definition from Chapter 7.

Definition An **(abelian) factor set** is a function $f: C \times C \to A$ such that, for all $x, y, z \in C$:

(i) $f(y, z) - f(x + y, z) + f(x, y + z) - f(x, y) = 0$;
(ii) $f(0, y) = 0 = f(x, 0)$;
(iii) $f(x, y) = f(y, x)$.

The main reason we used both additive and multiplicative notations in our previous discussion of extensions was that the quotient could operate nontrivially on the kernel: If $c \in C$, ca was a convenient notation for the element $l(c) + a - l(c)$. Since we are now considering only abelian extensions, as is ensured by condition (iii), $ca = a$ always; consequently, there is no need for separate notations. Furthermore, there is no longer a need to keep track of $\theta\colon C \to \mathrm{Aut}(A)$, for θ must be the trivial map here.

Definition $Z(C, A)$ is the additive group of all abelian factor sets.

We say that a function $g\colon C \times C \to A$ is a **coboundary** if there is a function $\alpha\colon C \to A$ with $\alpha(0) = 0$ for which

$$g(x, y) = \alpha(y) - \alpha(x + y) + \alpha(x).$$

Definition $B(C, A)$ is the set of all coboundaries.

As in Chapter 7, the reader may prove that $B(C, A)$ is a subgroup of $Z(C, A)$.

Definition $\mathrm{Ext}(C, A) = Z(C, A)/B(C, A)$.

When all groups under discussion are abelian, one uses the notation $\mathrm{Ext}(C, A)$ rather than $H^2(C, A)$. We now recall Exercise 7.36, which says that two extensions B and B' of A by C are equivalent in case there is a commutative diagram with exact rows

$$\begin{array}{ccccccccc} 0 & \to & A & \to & B & \to & C & \to & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \to & A & \to & B' & \to & C & \to & 0 \end{array}$$

where the outer downward maps are identities.

Theorem 10.10 *The set of equivalence classes of abelian extensions of A by C, $e(C, A)$, is a group isomorphic to* $\mathrm{Ext}(C, A)$*; the zero element is the class of the split extension.*

Proof We have only restated Theorem 7.15. ∎

Corollary 10.11 $\mathrm{Ext}(C, A) = 0$ *if and only if every short exact sequence* $0 \to A \to B \to C \to 0$ *splits.*

Corollary 10.12 *A group D is divisible if and only if* $\operatorname{Ext}(A, D) = 0$ *for every group A; a group F is free abelian if and only if* $\operatorname{Ext}(F, A) = 0$ *for every group A.*

Proof If D is divisible and $0 \to D \to E \to A \to 0$ is exact, then D is a direct summand of E, by Corollary 9.10, and so this sequence splits. Therefore $\operatorname{Ext}(A, D) = 0$.

Conversely, if every exact sequence $0 \to D \to E \to A \to 0$ splits, then D is divisible, by Corollary 9.24.

If F is free abelian and $0 \to A \to B \to F \to 0$ is exact, then Corollary 9.21 says this sequence must split; hence $\operatorname{Ext}(F, A) = 0$.

Conversely, choose an exact sequence $0 \to A \to B \to F \to 0$ in which B is free abelian (Corollary 9.19 says this is always possible). Since this sequence splits, F is a direct summand of B. Therefore F is free abelian, by Theorem 9.22. ▮

EXERCISES

10.16. If C is torsion-free and A is of bounded order, then $\operatorname{Ext}(C, A) = 0$.

10.17. If $\mathbf{R}$ denotes the additive group of real numbers, then $\operatorname{Ext}(\mathbf{R}, \sum \sigma(p)) \neq 0$. (HINT: Use Exercise 9.14.)

10.18. Prove that $\operatorname{Ext}(\sigma(p), \sigma(p)) \neq 0$.

10.19. If C is p-primary and A is q-primary, where $p \neq q$, then $\operatorname{Ext}(C, A) = 0$.

PULL-BACKS AND PUSH-OUTS

Let us give an explicit construction of the sum of two (classes of) extensions in $e(C, A)$. Rather than say that B is an extension of A by C, we now define an extension as a short exact sequence

$$0 \to A \to B \to C \to 0.$$

Suppose now that we have two extensions ($i = 1, 2$):

$$0 \longrightarrow A \xrightarrow{\lambda_i} B_i \xrightarrow{\pi_i} C \longrightarrow 0.$$

Suppose further that l_i are transversals and $f_i\colon C \times C \to A$ are the corresponding factor sets. What is the extension of A by C determined by the factor set $f_1 + f_2$?

Every $b_i \in B_i$ has the unique expression

$$b_i = a + l_i(c), \qquad a \in A, c \in C,$$

and we add according to the rule:

$$[a + l_i(c)] + [a' + l_i(c')] = a + a' + l_i(c + c') + f_i(c, c').$$

Let H be the subset of $B_1 \oplus B_2$ consisting of all elements (b_1, b_2) such that $\pi_1(b_1) = \pi_2(b_2)$. The reader may verify that H is a subgroup and that the following sequence is exact:

$$0 \to A \oplus A \xrightarrow{\lambda} H \xrightarrow{\mu} C \to 0,$$

where $\lambda(a, a') = (\lambda_1 a, \lambda_2 a')$ and $\mu(b_1, b_2) = \pi_1(b_1) = \pi_2(b_2)$. If we define a tranversal l by $l(c) = (l_1(c), l_2(c))$, then the usual computation defines a factor set F: $C \times C \to A \oplus A$ by $F(c, c') = (f_1(c, c'), f_2(c, c'))$. We must still produce an extension with a factor set that is the sum of the two coordinates.

Define ∇: $A \oplus A \to A$ by $\nabla(a, a') = a + a'$, and let S be the subgroup of $A \oplus H$ consisting of all elements of the form

$$(\nabla(a, a'), -\lambda(a, a')).$$

We shall prove that $(A \oplus H)/S$ is an extension of A by C; moreover, with respect to the transversal $\{l(c) + S\}$, the factor set of this extension is $f_1 + f_2$.

The complete construction thus consists of taking a subgroup of a direct sum and then taking a quotient of another direct sum. In the language of the preceding chapter, we perform dual constructions. Rather than supply more details now, we proceed to a systematic account of this construction.

Who would expect that a lemma about 10 groups and 13 homomorphisms could be of any use? Here is the "Five Lemma"; the method of proof is called "diagram-chasing".

Lemma 10.13 (The Five Lemma) *Consider the commutative diagram with exact rows:*

$$\begin{array}{ccccccccc} A_1 & \xrightarrow{f_1} & A_2 & \xrightarrow{f_2} & A_3 & \xrightarrow{f_3} & A_4 & \xrightarrow{f_4} & A_5 \\ \downarrow{\scriptstyle \alpha_1} & & \downarrow{\scriptstyle \alpha_2} & & \downarrow{\scriptstyle \alpha_3} & & \downarrow{\scriptstyle \alpha_4} & & \downarrow{\scriptstyle \alpha_5} \\ B_1 & \xrightarrow[g_1]{} & B_2 & \xrightarrow[g_2]{} & B_3 & \xrightarrow[g_3]{} & B_4 & \xrightarrow[g_4]{} & B_5 \end{array}$$

(i) *If α_2 and α_4 are onto and α_5 is one-to-one, then α_3 is onto.*

(ii) *If α_2 and α_4 are one-to-one and α_1 is onto, then α_3 is one-to-one.*

In particular, if α_1, α_2, α_4, and α_5 are isomorphisms, so is α_3.

REMARK The way to understand our proof is to chase elements around the diagram without pausing to give names to everything; we approximate this by using transparent notation, i.e., b_2 denotes an element of B_2, etc. The reader should observe that each step in the proof is automatic in the sense that there is only one reasonable way to proceed.

Proof We shall only prove (i), leaving the proof of the dual statement (ii) to the reader.

If $b_3 \in B_3$, we must find a_3 with $\alpha_3(a_3) = b_3$. Now $b_3 \to b_4 = g_3(b_3)$; since α_4 is onto, we may lift b_4 to a_4. Commutativity gives $a_5 f_4(a_4) = g_4\alpha_4(a_4) = g_4(b_4) = g_4 g_3(b_3)$, and this is 0 because the bottom row is exact. Since $\alpha_5 f_4(a_4) = 0$ and α_5 is one-to-one, $a_4 \in$ kernel f_4 = image f_3. Hence, there is $a_3 \in A_3$ with $f_3(a_3) = a_4$. Using commutativity in the next to last square, $g_3\alpha_3(a_3) = \alpha_4 f_3(a_3) = \alpha_4(a_4) = b_4 = g_3(b_3)$. Conclusion: $b_3 - \alpha_3(a_3)$ is in kernel g_3 = image g_2. Therefore, $b_3 - \alpha_3(a_3) = g_2(b_2)$ for some b_2. Since α_2 is onto, we may lift b_2 to a_2. Using commutativity once more, $\alpha_3 f_2(a_2) = g_2\alpha_2(a_2) = g_2(b_2) = b_3 - \alpha_3(a_3)$. Thus, $b_3 = \alpha_3(f_2(a_2) + a_3)$ and α_3 is onto. ∎

Corollary 10.14 *If B and B' are extensions of A by C, then B is equivalent to B' if and only if there is a commutative diagram with exact rows:*

$$\begin{array}{ccccccccc} 0 & \to & A & \to & B & \to & C & \to & 0 \\ & & {\scriptstyle 1_A}\downarrow & & \downarrow{\scriptstyle \gamma} & & \downarrow{\scriptstyle 1_C} & & \\ 0 & \to & A & \to & B' & \to & C & \to & 0 \end{array}$$

Proof The five lemma tells us that γ is automatically an isomorphism. ∎

A weaker form of the five lemma says that if the four outer maps are isomorphisms and if there is a middle map α_3 such that all commutes, then α_3 is an isomorphism. We remark that a map α_3 need not exist, for consider the diagram with exact rows:

$$\begin{array}{ccccccccc} 0 & \to & \sigma(p) & \to & \sigma(p) \oplus \sigma(p) & \to & \sigma(p) & \to & 0 \\ & & \downarrow & & & & \downarrow & & \\ 0 & \to & \sigma(p) & \longrightarrow & \sigma(p^2) & \longrightarrow & \sigma(p) & \to & 0 \end{array}$$

where the outer maps are identities. One cannot insert a homomorphism in the middle making the resulting diagram commute, lest it be an isomorphism between the two nonisomorphic middle groups.

EXERCISES

**10.20. Consider the commutative diagram with exact rows:

$$\begin{array}{ccccccc} A_1 & \to & A_2 & \to & A_3 & \to & 0 \\ {\scriptstyle \alpha_1}\downarrow & & {\scriptstyle \alpha_2}\downarrow & & & & \\ B_1 & \to & B_2 & \to & B_3 & \to & 0 \end{array}$$

Prove there is a homomorphism from A_3 to B_3 making the augmented diagram commute; this map is an isomorphism if α_1 and α_2 are.

**10.21. Consider the diagram

$$\begin{array}{ccccccc} 0 & \to & R & \to & F & \to & C & \to & 0 \\ & & & & & & {\scriptstyle 1_C}\downarrow & & \\ & & E & \to & B & \to & C & \to & 0 \end{array}$$

where the rows are exact and F is free. There exist maps $\alpha\colon F \to B$ and $\beta\colon R \to E$ such that the resulting diagram commutes.

**10.22. Consider the diagram

$$\begin{array}{ccccccc} 0 & \to & A & \to & B & \to & E \\ & & {\scriptstyle 1_A}\downarrow & & & & \\ 0 & \to & A & \to & D & \to & D' & \to & 0 \end{array}$$

where the rows are exact and D is divisible. There exist maps $\alpha\colon B \to D$ and $\beta\colon E \to D'$ such that the resulting diagram commutes.

Definition Consider the extensions

$$E\colon 0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0$$

and

$$E'\colon 0 \to A' \xrightarrow{\lambda'} B' \xrightarrow{\mu'} C' \to 0;$$

a **map** from E to E' is an ordered triple (α, β, γ) of homomorphisms such that the following diagram commutes:

$$\begin{array}{ccccccccc} E\colon 0 & \to & A & \xrightarrow{\lambda} & B & \xrightarrow{\mu} & C & \to & 0 \\ & & {\scriptstyle \alpha}\downarrow & & {\scriptstyle \beta}\downarrow & & {\scriptstyle \gamma}\downarrow & & \\ E'\colon 0 & \to & A' & \xrightarrow{\lambda'} & B' & \xrightarrow{\mu'} & C' & \to & 0 \end{array}$$

Our notation is $(\alpha, \beta, \gamma)\colon E \to E'$.

One composes maps of extensions "coordinatewise". The definition of equivalence of extensions E and E' states that there is a map $(1_A, \beta, 1_C)$: $E \to E'$; Corollary 10.14 implies that this relation is an equivalence relation. The equivalence class of an extension E is denoted $[E]$.

There are two dual operations on extensions. We now prepare the first one.

By a **solution** of the diagram

(*)
$$\begin{array}{ccc} & & C' \\ & & \downarrow{\scriptstyle \gamma} \\ B & \xrightarrow{\mu} & C \end{array}$$

we mean a group B' and maps μ' and γ' such that

$$\begin{array}{ccc} B' & \xrightarrow{\mu'} & C' \\ {\scriptstyle \gamma'}\downarrow & & \downarrow{\scriptstyle \gamma} \\ B & \xrightarrow{\mu} & C \end{array}$$

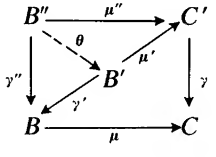
commutes.

A solution of a diagram need not be unique: If X is any group and β: $X \to B'$ any homomorphism, then

$$\begin{array}{ccc} X & \xrightarrow{\beta\mu'} & C' \\ {\scriptstyle \beta\gamma'}\downarrow & & \downarrow{\scriptstyle \gamma} \\ B & \xrightarrow[\mu]{} & C \end{array}$$

is also a solution.

Definition A solution of the diagram (*) is a **pull-back** in case, given any other solution (B'', μ'', γ'') of (*), there is a unique map θ: $B'' \to B'$ making the following diagram commute:

(**)


Lemma 10.15 *A pull-back exists for the diagram* (*).

Proof Let B' be the subgroup of $B \oplus C'$ consisting of all (b, c') such that $\mu b = \gamma c'$, and define μ' and γ' as projections. It is trivial to verify that we have defined a solution of (*).

Suppose now that we have a second solution (diagram (**)). Define $\theta\colon B'' \to B'$ as follows: If $x \in B''$, then $\theta x = (\gamma'' x, \mu'' x)$. One checks quickly that (**) commutes and this is the only θ that works. ∎

Theorem 10.16 *Let E be an extension of A by C and let $\gamma\colon C' \to C$ be a homomorphism. There is an extension E' of A by C' and a map $(1_A, \gamma', \gamma)\colon E' \to E$.*

Proof We are given the following diagram, which we must complete:

$$\begin{array}{ccccccccc} & & A & & C' & & \\ & & {\scriptstyle 1_A}\downarrow & & \downarrow{\scriptstyle \gamma} & & \\ 0 & \to & A & \xrightarrow{\lambda} & B & \xrightarrow{\mu} & C & \to & 0 \end{array}$$

Let (B', μ', γ') be the pull-back obtained in Lemma 10.15 and define $\lambda'\colon A \to B'$ by $\lambda' a = (\lambda a, 0)$. It is easy to check that

$$E'\colon 0 \to A \xrightarrow{\lambda'} B' \xrightarrow{\mu'} C' \to 0$$

is an extension and that $(1_A, \gamma', \gamma)\colon E' \to E$. ∎

The extension E' just constructed is denoted $E\gamma$.

Notation If $\gamma\colon C' \to C$ is a function, then we define

$$\gamma \times \gamma\colon C' \times C' \to C \times C \text{ by } \gamma \times \gamma(c_1', c_2') = (\gamma c_1', \gamma c_2').$$

If $\lambda\colon A \to B$ and $\lambda'\colon A' \to B'$ are homomorphisms, we define

$$\lambda \oplus \lambda'\colon A \oplus A' \to B \oplus B' \text{ by } \lambda \oplus \lambda'(a, a') = (\lambda a, \lambda' a').$$

It is useful to do a bit of bookkeeping.

Lemma 10.17 *Let E be an extension of A by C and let $\gamma\colon C' \to C$ be a homomorphism. If $f\colon C \times C \to A$ is a factor set of E, then $f(\gamma \times \gamma)$ is a factor set of $E\gamma$.*

Proof We use the notation of the preceding theorem. Let l be a transversal determining f. Define a transversal L by $L(c') = (l(\gamma c'), c')$. The reader may now compute that the corresponding factor set is $F = f(\gamma \times \gamma)$, i.e., $F(c_1', c_2') = f(\gamma c_1', \gamma c_2')$. ∎

Corollary 10.18 *Let E and E'' be extensions of A by C, and let $\gamma\colon C' \to C$. If $[E] = [E'']$, then $[E\gamma] = [E''\gamma]$.*

Proof Let f and f'' be factor sets of E and E'', respectively; then $f - f'' \in B(C, A)$. By Lemma 10.17, $f(\gamma \times \gamma)$ and $f''(\gamma \times \gamma)$ are factor sets of $E\gamma$ and $E''\gamma$, respectively, and the reader may check that $f(\gamma \times \gamma) - f''(\gamma \times \gamma) \in B(C', A)$. Hence, $[E\gamma] = [E''\gamma]$. ∎

Corollary 10.19 *Let E be an extension of A by C and let $\gamma: C' \to C$ and $\gamma_1: C'' \to C'$ be homomorphisms. Then $[(E\gamma)\gamma_1] = [E(\gamma\gamma_1)]$.*

Proof If f is a factor set of E, then $[f \circ (\gamma \times \gamma)] \circ (\gamma_1 \times \gamma_1)$ is a factor set of $(E\gamma)\gamma_1$ and $f \circ [(\gamma \times \gamma) \circ (\gamma_1 \times \gamma_1)]$ is a factor set of $E(\gamma\gamma_1)$. Since composition of functions is associative, $[(E\gamma)\gamma_1] = [E(\gamma\gamma_1)]$. ∎

We now prepare our second construction. By a **solution** of the diagram

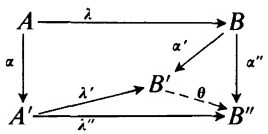
(♯)
$$\begin{array}{ccc} A & \xrightarrow{\lambda} & B \\ {\scriptstyle\alpha}\downarrow & & \\ A' & & \end{array}$$

we mean a group B' and maps λ' and α' such that

$$\begin{array}{ccc} A & \xrightarrow{\lambda} & A \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\alpha'} \\ A' & \xrightarrow{\lambda'} & B' \end{array}$$

commutes.

Definition A solution of the diagram (♯) is a **push-out** in case, given any other solution $(B'', \lambda'', \alpha'')$ of (♯), there is a unique map $\theta: B' \to B''$ making the following diagram commute:

(♯♯)


Lemma 10.20 *A push-out exists for the diagram* (♯).

Proof Since the dual of a subgroup is a quotient, one should expect the push-out to be a quotient of $A' \oplus B$. If S is the subgroup of $A' \oplus B$ consisting of all $(\alpha a, -\lambda a)$, where $a \in A$, then define $B' = (A' \oplus B)/S$. Define $\lambda': A' \to B'$ by $\lambda' a' = (a', 0) + S$; define $\alpha': B \to B'$ by

$\alpha' b = (0, b) + S$. It is a trivial matter to verify that we have defined a solution of (♯).

Suppose now that we have a second solution (diagram (♯♯)). Define $\theta: B' \to B''$ by $\theta((a', b) + S) = \lambda'' a' + \alpha'' b$. One checks easily that θ is well defined, that (♯♯) commutes, and that this is the only θ that works. ∎

Theorem 10.21 *Let E be an extension of A by C and let $\alpha: A \to A'$ be a homomorphism. There is an extension E' of A' by C and a map $(\alpha, \alpha', 1_C): E \to E'$.*

Proof We are given the following diagram to complete:

$$\begin{array}{ccccccccc} 0 & \to & A & \xrightarrow{\lambda} & B & \xrightarrow{\mu} & C & \to & 0 \\ & & {\scriptstyle\alpha}\downarrow & & & & \downarrow{\scriptstyle 1_C} & & \\ & & A' & & & & C & & \end{array}$$

Let (B', λ', α') be a push-out, and define $\mu': B' \to C$ by $\mu'((a', b) + S) = \mu b$. It is a simple matter to verify that μ' is well defined,

$$E': 0 \to A' \xrightarrow{\lambda'} B' \xrightarrow{\mu'} C' \to 0$$

is an extension, and $(\alpha, \alpha', 1_C): E \to E'$. ∎

We denote the extension E' just constructed by αE.

Lemma 10.22 *Let E be an extension of A by C and let $\alpha: A \to A'$ be a homomorphism. If $f: C \times C \to A$ is a factor set of E, then αf is a factor set of αE.*

Proof We use the notation of Theorem 10.21. If l is a transversal determining f, define a transversal L by $L(c) = (0, l(c)) + S$. Now the factor set F determined by this choice is $F(c_1, c_2) = L(c_1) + L(c_2) - L(c_1 + c_2) = (0, f(c_1, c_2)) + S = (\alpha f(c_1, c_2), 0) + S = \lambda' \alpha f(c_1, c_2)$. We are done, for λ' is only the inclusion map. ∎

Corollary 10.23 *If $[E] = [E'']$, then $[\alpha E] = [\alpha E'']$.*

Corollary 10.24 *If E is an extension of A by C and if $\alpha: A \to A'$ and $\alpha_1: A' \to A''$ are homomorphisms, then $[\alpha_1(\alpha E)] = [(\alpha_1 \alpha)E]$.*

Corollary 10.25 *Let E be an extension of A by C and let $\alpha: A \to A'$ and $\gamma: C' \to C$ be homomorphisms. Then $[(\alpha E)\gamma] = [\alpha(E\gamma)]$.*

Proof If f is a factor set of E, then $(\alpha f)\gamma \times \gamma$ is a factor set of $(\alpha E)\gamma$ and $\alpha(f(\gamma \times \gamma))$ is a factor set of $\alpha(E\gamma)$. ∎

We have demonstrated that homomorphisms operate on extensions and that all possible associativity laws hold.

Definition Consider the extensions

$$E\colon 0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0$$

and

$$E'\colon 0 \to A' \xrightarrow{\lambda'} B' \xrightarrow{\mu'} C' \to 0.$$

Their **direct sum** is the extension

$$E \oplus E'\colon 0 \longrightarrow A \oplus A' \xrightarrow{\lambda \oplus \lambda'} B \oplus B' \xrightarrow{\mu \oplus \mu'} C \oplus C' \longrightarrow 0.$$

Notation The **diagonal map** $\Delta\colon C \to C \oplus C$ is defined by $\Delta(c) = (c, c)$; the **codiagonal map** $\nabla\colon A \oplus A \to A$ is defined by $\nabla(a, a') = a + a'$.

Definition If E and E' are extensions of A by C, their sum $E + E'$ is the extension $\nabla(E \oplus E')\Delta$.

The reader should observe that $E + E'$ is an extension of A by C (whose equivalence class is well defined, by Corollary 10.25).

EXERCISES

10.23. If E and E' are extensions with factor sets f and f', respectively, then $f \oplus f'$ is a factor set of $E \oplus E'$.

10.24. If $[E] = [E']$ and $[E_0] = [E'_0]$, then $[E \oplus E_0] = [E' \oplus E'_0]$.

10.25. Let E be an extension of A by C, E' an extension of A' by C', and $\gamma\colon C_1 \to C$, $\gamma'\colon C_2 \to C'$ homomorphisms. Then

$$[(E \oplus E')(\gamma \oplus \gamma')] = [E\gamma \oplus E'\gamma'].$$

A similar equation holds on the left.

***10.26.** If $[E] = [E']$ and $[E_0] = [E'_0]$, then $[E + E_0] = [E' + E'_0]$.

10.27. $[E(\gamma + \gamma_1)] = [E\gamma + E\gamma_1]$; similarly on the left.

****10.28.** $[(E + E')\gamma] = [E\gamma + E'\gamma]$; similarly on the left.

The following result shows that our constructions do give an extension corresponding to the sum of two factor sets.

Lemma 10.26 *If f is a factor set, let E_f be an extension it determines (so that $\varphi(f + B(C, A)) = [E_f]$ defines the one-to-one correspondence of Theorem* 10.10). *Then $[E_{f+f'}] = [E_f + E_{f'}]$.*

Proof By definition, $E_{f+f'}$ has $f + f'$ as a factor set; $E_f + E_{f'}$ has $\nabla(f \oplus f')\Delta$ as a factor set. But these are the same:

$$\begin{aligned}\nabla(f \oplus f')\,\Delta(x, y) &= \nabla(f \oplus f')(x, y, x, y)\\ &= \nabla(f(x, y), f'(x, y))\\ &= f(x, y) + f'(x, y)\\ &= (f + f')(x, y). \quad \blacksquare\end{aligned}$$

Definition If $[E]$ and $[E'] \in e(C, A)$, their **Baer sum** is $[E] + [E'] = [E + E']$.

By Exercise 10.26, the Baer sum is a well-defined binary operation on $e(C, A)$.

Theorem 10.27 *$e(C, A)$ is an abelian group under Baer sum, and φ: $\mathrm{Ext}(C, A) \to e(C, A)$ defined by $\varphi(f + B(C, A)) = [E_f]$ is an isomorphism.*

Proof We have known since Chapter 7 that φ is a one-to-one correspondence. Lemma 10.26 and Exercise 1.19 complete the proof. $\blacksquare$

THE EXT FUNCTORS

We had an ulterior motive in examining the addition of extensions, for we are now in a position to prove that Ext defines functors that are closely related to the Hom functors.

Theorem 10.28 *If G is a fixed group, then* $\mathrm{Ext}(G, \ )$ *is a covariant functor.*

Proof If $\alpha: A \to A'$, we define a homomorphism

$$\alpha_*: \mathrm{Ext}(G, A) \to \mathrm{Ext}(G, A')$$

by

$$\alpha_*(f + B(G, A)) = \alpha f + B(G, A').$$

It is mechanical to prove that α_* is a well-defined homomorphism, and all the axioms of the definition of functor are satisfied. ∎

Theorem 10.29 *If K is a fixed group, then* $\mathrm{Ext}(\ , K)$ *is a contravariant functor.*

Proof If $\alpha: A \to A'$, define $\alpha^*: \mathrm{Ext}(A', K) \to \mathrm{Ext}(A, K)$ by $\alpha^*(f + B(A', K)) = f(\alpha \times \alpha) + B(A, K)$. ∎

We have another way of looking at extensions, $e(C, A)$, and this also gives rise to functors. If $\alpha: A \to A'$, define a function $\alpha_\#: e(C, A) \to e(C, A')$ by

$$\alpha_\#([E]) = [\alpha E];$$

if $\gamma: C \to C'$, define a function $\gamma^\#: e(C', A) \to e(C, A)$ by

$$\gamma^\#([E]) = [E\gamma].$$

The following lemma will imply that these functions are homomorphisms (this also follows from Exercise 10.28).

Lemma 10.30[1] *If $\alpha: A \to A'$, then the following diagram commutes for every G:*

$$\begin{array}{ccc} \mathrm{Ext}(G, A) & \xrightarrow{\alpha_*} & \mathrm{Ext}(G, A') \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ e(G, A) & \xrightarrow{\alpha_\#} & e(G, A') \end{array}$$

A dual theorem holds if we fix the second variable.

Proof We must show that $\varphi\alpha_* = \alpha_\#\varphi$. If $f \in Z(G, A)$, then $\varphi\alpha_*(f + B) = \varphi(\alpha f + B') = [E_{\alpha f}]$, i.e., the class of an extension with factor set αf. On the other hand, $\alpha_\#\varphi(f + B) = \alpha_\#[E_f] = [\alpha E_f]$. By Lemma 10.22, αf is a factor set of αE_f, and so $[\alpha E_f] = [E_{\alpha f}]$. ∎

[1] This lemma says the functors $\mathrm{Ext}(G, \ )$ and $e(G, \ )$ are **naturally equivalent.**

Theorem 10.31 *The functions $\alpha_\#$ and $\gamma^\#$ are homomorphisms.*

Proof By the lemma, $\alpha_\# = \varphi\alpha_*\varphi^{-1}$, and so is the composite of homomorphisms. ∎

We now connect Hom to Ext.

Definition Let f be a factor set of the extension

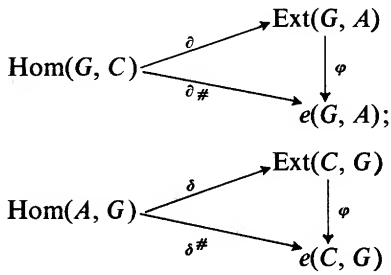
$$E\colon 0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0.$$

If G is a group, we define the **connecting homomorphisms**:

$$\partial\colon \mathrm{Hom}(G, C) \to \mathrm{Ext}(G, A) \quad \text{by} \quad \partial(h) = f(h \times h) + B(G, A);$$
$$\delta\colon \mathrm{Hom}(A, G) \to \mathrm{Ext}(C, G) \quad \text{by} \quad \delta(h') = h'f + B(C, G).$$

EXERCISE

10.29. Define $\partial_\#\colon \mathrm{Hom}(G, C) \to e(G, A)$ by $\partial_\#(h) = [Eh]$, and $\delta^\#\colon \mathrm{Hom}(A, G) \to e(C, G)$ by $\delta^\#(h') = [h'E]$. Prove the following diagrams commute:



Two lemmas are needed before the main tool for calculating Ext can be presented. The proofs are dual, so the second proof is left as an exercise.

Lemma 10.32 *If $0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0$ is exact, then, for every group G, the map $\lambda^\#\colon e(B, G) \to e(A, G)$ is onto.*

Proof Let $0 \to G \xrightarrow{\alpha} E \xrightarrow{\beta} A \to 0$ represent an element of $e(A, G)$. We now hook this sequence to our given sequence to get the following exact sequence:

$$0 \to G \xrightarrow{\alpha} E \xrightarrow{\lambda\beta} B \xrightarrow{\mu} C \to 0.$$

There exists an exact sequence

$$0 \to R \to F \to C \to 0,$$

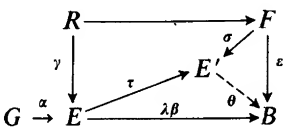
where F is free. By Exercise 10.21 there are maps such that the following diagram commutes:

$$\begin{array}{ccccccccc} 0 & \to & R & \to & F & \to & C & \to & 0 \\ & & \downarrow{\scriptstyle \gamma} & & \downarrow{\scriptstyle \varepsilon} & & \downarrow{\scriptstyle 1_C} & & \\ 0 & \to & G & \to & E & \xrightarrow{\lambda\beta} & B & \to & C \to 0. \end{array}$$

If we focus on the first square in this diagram, then we see that $(B, \lambda\beta, \varepsilon)$ is a solution of the diagram

$$\begin{array}{ccc} R & \to & F \\ \downarrow & & \\ E & & \end{array}$$

If (E', τ, σ) is the push-out for this diagram, then there is a map $\theta\colon E' \to B$ with



commuting. The reader may now verify that

$$0 \to G \xrightarrow{\tau\alpha} E' \xrightarrow{\theta} B \to 0$$

is an extension, and that

$$\begin{array}{ccccccccc} 0 & \to & G & \to & E & \to & A & \to & 0 \\ & & \downarrow{\scriptstyle 1_G} & & \downarrow{\scriptstyle \tau} & & \downarrow{\scriptstyle \lambda} & & \\ 0 & \to & G & \to & E' & \to & B & \to & 0 \end{array}$$

commutes. In other words, $\lambda^{\#}$ sends the class of the bottom extension into the class of the top extension. ∎

Lemma 10.33 *If* $0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0$ *is exact, then for every group* G *the map* $\mu_{\#}\colon e(G, B) \to e(G, C)$ *is onto.*

Proof Dual to the preceding proof, using pull-backs instead of push-outs and Exercise 10.22 instead of Exercise 10.21. ∎

Theorem 10.34 *Let* $0 \to A \xrightarrow{\lambda} B \xrightarrow{\mu} C \to 0$ *be exact. For any group* G, *the following two sequences are exact:*

$$\text{(i)}\quad 0 \longrightarrow \operatorname{Hom}(G, A) \xrightarrow{T(\lambda)} \operatorname{Hom}(G, B) \xrightarrow{T(\mu)} \operatorname{Hom}(G, C) \xrightarrow{\partial} \operatorname{Ext}(G, A) \xrightarrow{\lambda^*} \operatorname{Ext}(G, B) \xrightarrow{\mu^*} \operatorname{Ext}(G, C) \longrightarrow 0;$$

$$\text{(ii)}\quad 0 \longrightarrow \operatorname{Hom}(C, G) \xrightarrow{S(\mu)} \operatorname{Hom}(B, G) \xrightarrow{S(\lambda)} \operatorname{Hom}(A, G) \xrightarrow{\delta} \operatorname{Ext}(C, G) \xrightarrow{\mu^*} \operatorname{Ext}(B, G) \xrightarrow{\lambda^*} \operatorname{Ext}(A, G) \longrightarrow 0.$$

Proof The difficult portions are the proofs that μ_* and λ^* are onto, and they follow easily from Lemmas 10.32 and 10.33 and Lemma 10.30.

The other steps of the proof are utterly uninspiring, proceeding inexorably in the manner of the proof of Theorem 10.3. ∎

The groups Ext thus repair the exactness we may have lost by applying Hom. Consider the question: If A is a subgroup of B, when can a homomorphism $h: A \to G$ be extended to a homomorphism of B into G? The exact sequence answers this question. Let $0 \to A \xrightarrow{i} B \to B/A \to 0$ be exact (where i is the inclusion). Then we have exactness of

$$\operatorname{Hom}(B, G) \xrightarrow{S(i)} \operatorname{Hom}(A; G) \xrightarrow{\delta} \operatorname{Ext}(B/A, G).$$

Definition The element $\delta(h)$ is the **obstruction** of h.

Corollary 10.35 *Let* A *be a subgroup of* B *and* $h: A \to G$ *a homomorphism. Then* h *can be extended to* B *if and only if its obstruction is* 0.

Proof If h can be extended to B, there is a map $h': B \to G$ such that $h'i = h$. Hence, $h \in$ image $S(i) =$ kernel δ, so that $\delta(h) = 0$.

If $\delta(h) = 0$, then $h \in$ kernel $\delta =$ image $S(i)$, so that there is an $h' \in \operatorname{Hom}(B, G)$ with $S(i)(h') = h$, i.e., $h'i = h$, as desired. ∎

Corollary 10.36 *If* $\operatorname{Ext}(B/A, G) = 0$, *every map* $h: A \to G$ *can be extended to* B.

We now prove that Ext behaves very much like Hom.

Theorem 10.37 *Let* $\{A_i: i \in I\}$ *be a family of groups. For any* K,

$$\operatorname{Ext}(\textstyle\sum A_i, K) \cong \prod \operatorname{Ext}(A_i, K).$$

Proof For each i, let $0 \to R_i \to F_i \to A_i \to 0$ be exact, where F_i is free. Consider the commutative diagram with exact rows:

$$\begin{array}{ccccccc} \mathrm{Hom}(\sum F_i, K) & \to & \mathrm{Hom}(\sum R_i, K) & \to & \mathrm{Ext}(\sum A_i, K) & \to & \mathrm{Ext}(\sum F_i, K) \\ \downarrow & & \downarrow & & & & \\ \prod \mathrm{Hom}(F_i, K) & \to & \prod \mathrm{Hom}(R_i, K) & \to & \prod \mathrm{Ext}(A_i, K) & \to & \prod \mathrm{Ext}(F_i, K) \end{array}$$

where the downward maps are the isomorphisms of Theorem 10.7. By Corollary 10.12 the terms on the far right are 0, so that Exercise 10.20 yields the desired result. ∎

Theorem 10.38 *Let* $\{A_i : i \in I\}$ *be a family of groups. For any* G,

$$\mathrm{Ext}(G, \prod A_i) \cong \prod \mathrm{Ext}(G, A_i).$$

Proof Left to the reader. ∎

Theorem 10.39 *For any group* A, $\mathrm{Ext}(\sigma(n), A) \cong A/nA$.

Proof Consider the exact sequence

$$0 \to Z \xrightarrow{\mu} Z \to \sigma(n) \to 0$$

where μ is multiplication by n. We now have exactness of

$$\mathrm{Hom}(Z, A) \xrightarrow{S(\mu)} \mathrm{Hom}(Z, A) \longrightarrow \mathrm{Ext}(\sigma(n), A) \longrightarrow \mathrm{Ext}(Z, A) = 0.$$

Now $\mathrm{Hom}(Z, A) \cong A$ and $S(\mu)$ is multiplication by n, so that our result follows. ∎

It follows from Theorem 10.39 that $\mathrm{Ext}(\sigma(p), \sigma(p)) \cong \sigma(p)$, where p is a prime, so that this Ext has p elements. On the other hand, if $0 \to \sigma(p) \to E \to \sigma(p) \to 0$ is exact, then E has order p^2 so that $E \cong \sigma(p^2)$ or $E \cong \sigma(p) \oplus \sigma(p)$. Therefore, if p is an odd prime, we have an example of inequivalent extensions of A by C with isomorphic middle terms.

The number of nonisomorphic middle groups B (where $0 \to A \to B \to C \to 0$ is exact) is at most $|\mathrm{Ext}(C, A)|$. Our example above shows that this inequality can be strict.

Example 5 If A is torsion-free, then $\mathrm{Ext}(A, G)$ is divisible for any group G.

Since A is torsion-free, for each $n > 0$ there is an exact sequence

$$0 \to A \xrightarrow{n} A \to A/nA \to 0,$$

where the first nontrivial map is multiplication by n. This gives rise to an exact sequence

$$\cdots \to \operatorname{Ext}(A, G) \xrightarrow{n} \operatorname{Ext}(A, G) \to 0,$$

so that multiplication by n maps $\operatorname{Ext}(A, G)$ onto itself. This says that $\operatorname{Ext}(A, G)$ is divisible.

Example 6 If $mG = 0$, then $\operatorname{Ext}(A, G) = 0$ for every torsion-free group A.

If $\mu: G \to G$ is multiplication by m, then $\mu_*: \operatorname{Ext}(A, G) \to \operatorname{Ext}(A, G)$ is also multiplication by m. Since $mG = 0$, μ is the zero map; thus μ_* is also the zero map. Therefore

$$\mu_* \operatorname{Ext}(A, G) = m \operatorname{Ext}(A, G) = 0.$$

On the other hand, we have just seen that $\operatorname{Ext}(A, G)$ is divisible, for A is torsion-free. It follows that $\operatorname{Ext}(A, G) = 0$.

We have given a proof by homological algebra of Corollary 9.37: If B is a group with tB of bounded order, then tB is a direct summand of B. The group B/tB is torsion-free, and by Corollary 10.11, $\operatorname{Ext}(B/tB, tB) = 0$ implies that the sequence

$$0 \to tB \to B \to B/tB \to 0$$

splits.

EXERCISES

10.30. Prove that $\operatorname{Ext}(V, G)$ is torsion-free and divisible whenever V is a vector space over $\mathbf{Q}$.

10.31. Prove that $\operatorname{Ext}(\mathbf{Q}, Z) \neq 0$.

10.32. Prove that

$$\operatorname{Ext}\left(\mathbf{Q}, \sum_{n=1}^{\infty} \sigma(p^n)\right) \neq 0.$$

Conclude that there exists a (countable) group G whose torsion subgroup $tG \cong \sum \sigma(p^n)$ and such that tG is not a summand of G.

10.33. Let $0 \to R \to F \to A \to 0$, where F is free, be exact. Prove that $\operatorname{Ext}(A, G) \cong \operatorname{Hom}(R, G)/\text{image } \operatorname{Hom}(F, G)$. (This is often taken as the definition of $\operatorname{Ext}(A, G)$.)

10.34. Prove that $\operatorname{Ext}(A, G) = 0$ for every torsion-free group A if and only if $\operatorname{Ext}(\mathbf{Q}, G) = 0$.

10.35. (**Nunke**) If $\operatorname{Hom}(A, Z) = 0$ and $\operatorname{Ext}(A, Z) = 0$, then $A = 0$.

10.36. If A is a torsion-free group of rank 1 such that $\text{Ext}(A, Z) = 0$, then $A \cong Z$.

10.37. A group G is divisible if and only if $\text{Ext}(\mathbf{Q}/Z, G) = 0$.

10.38. Prove that an abelian group A is free if and only if $\text{Ext}(A, F) = 0$ for every free abelian group F.

10.39. If G is a torsion group with character group G^*, then $G^* \cong \text{Ext}(G, Z)$.

10.40. Let W be a group with $\text{Ext}(W, Z) = 0$. Prove that W is torsion-free. If W has finite rank, prove that W is free abelian.

If K is an abelian group and G is a not necessarily abelian extension of K by a group Q, then we remarked in Chapter 7 that K is a module over a ring ZQ. The constructions of this chapter involved Z-modules (abelian groups); the more general theory involves ZQ-modules. The main tool of the general theory is the existence of a sequence of functors $\text{Ext}^n(\ ,\ )$ and a long exact sequence beginning with a triple of Hom, then a triple of Ext^1, a triple of Ext^2, etc. If one considers the group Z as a ZQ-module (by defining scalar multiplication suitably), then $H_\theta^2(Q, K) \cong \text{Ext}^2(Z, K)$. We refer the reader to the books of Mac Lane, Gruenberg, Babakhanian, and Rotman for further details.

CHAPTER

11

Free Groups and Free Products

GENERATORS AND RELATIONS

In this chapter, we shall consider several special classes of infinite groups.

Recall first that a free abelian group F on generators X has the following basic[1] property: If G is an arbitrary abelian group, then any function $f: X \to G$ may be extended to a homomorphism of F into G. An immediate consequence of this property is Corollary 9.19: Every abelian group is a quotient of a free abelian group. The significance of this remark is that we now have a convenient way of describing abelian groups.

Definition An abelian group G has **generators** $X = \{x_k : k \in K\}$ and **relations** $\{r_j = 0 : j \in J\}$ in case $G \cong F/R$, where F is free abelian on X, and R is the subgroup generated by $\{r_j : j \in J\}$.

In each of the following examples, we present abelian groups by generators and relations.

Example 1 $G = \sigma(6)$ has generator x and relation $6x = 0$.

Example 2 $G = \sigma(6)$ has generators $\{x, y\}$ and relations $\{2x = 0, 3y = 0\}$.

[1] A pun is intended; if V is a vector space with basis X and if W is another vector space, then any function $f: X \to W$ has an extension to a linear transformation of V into W.

Example 3 $G = \sigma(p^\infty)$ has generators

$$\{a_1, a_2, \cdots, a_n, \cdots\}$$

and relations

$$\{pa_1 = 0, pa_n = a_{n-1} \text{ if } n > 1\}.$$

Example 4 If G is free abelian on $\{x_k : k \in K\}$, then G has generators $\{x_k : k \in K\}$ and no relations (recall that $\{0\}$ is the subgroup generated by the empty set). The etymology of the term *free* is apparent now.

We have seen that we can describe an existing group by generators and relations. We can also use generators and relations to construct a group having prescribed properties. For example, is there a reduced p-primary abelian group G such that

$$\bigcap_{n=1}^{\infty} p^n G \neq \{0\}?$$

Construct such a group G as the abelian group with generators

$$a, b_1, b_2, \cdots, b_n, \cdots$$

and relations

$$pa = 0 \quad \text{and} \quad p^n b_n = a \qquad \text{for all } n.$$

It is easy to see that G is a p-primary abelian group and that $a \in p^n G$ for all n. We shall prove that $a \neq 0$, leaving the proof that G is reduced, using Exercise 9.37, to the reader.

Let F be the free abelian group on generators

$$a, b_1, b_2, \cdots, b_n, \cdots$$

and let R be the subgroup of F generated by

$$\{pa, p^n b_n - a, \quad n \geq 1\}.$$

We must prove that $a \notin R$. If, on the contrary, $a \in R$, then

$$a = mpa + \sum m_n(p^n b_n - a)$$

for integers m and m_n. Collecting terms gives

$$(1 - mp + \sum m_n)a = \sum m_n p^n b_n.$$

Since $\{a, b_1, b_2, \cdots\}$ is a basis, each $m_n p^n = 0$ and $1 - mp + \sum m_n = 0$. From these equations, it follows that each $m_n = 0$, and so $1 = mp$, a contradiction.

The reader should note that an abelian group G defined by generators and relations is the largest such group in the sense that any other group generated by elements subject to these defining relations is a quotient of G.

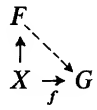
For example, let G be the abelian group with generators $\{x, y\}$ and relations $\{2x = 0, 4y = 0\}$. The reader may check that

$$G = \sigma(2) \oplus \sigma(4).$$

On the other hand, if $H = \sigma(2) \oplus \sigma(2)$, then H is also generated by two elements that satisfy the defining relations of G (of course they satisfy an additional relation as well).

We wish to extend the idea of generators and relations from abelian groups to arbitrary groups. In order to do this, we need the nonabelian analog of free abelian groups, and so we abstract the basic property of a free set of generators.

Definition Let X be a set and F a group containing X; F is **free** on X if for every group G, every function $f: X \to G$ has a unique extension to a homomorphism of F into G. One calls X a **basis** of F.



We shall prove later that the uniqueness of the extension is equivalent to saying that the set X generates F.

The first question is whether there are any free groups.

Theorem 11.1 *If X is a set, then there exists a group F that is free on X.*

Proof Let X' be a set disjoint from X and in one-to-one correspondence with it; we denote this correspondence

$$x \leftrightarrow x^{-1}.$$

Let X'' be a set disjoint from $X \cup X'$ that contains only one element which we denote "1". Call $X \cup X' \cup X''$ the *alphabet*, and call its elements *letters*. Let S be the set of all sequences of letters. A **word** in X is a sequence

$$w = (a_1, a_2, \cdots) \in S$$

such that all coordinates are 1 from some point on, i.e., there is an integer n such that $a_k = 1$ for all $k \geq n$. In particular, the constant sequence

$$(1, 1, 1, \cdots)$$

is a word; it is called the **empty word**, and is also denoted 1. A **reduced word** in X is a word in X that satisfies the extra conditions:

(i) x and x^{-1} are never adjacent;
(ii) If $a_m = 1$ for some m, then $a_k = 1$ for all $k > m$.

In particular, the empty word is a reduced word. Since words contain only a finite number of letters before they become constant, we use the more economical (and suggestive) notation:

$$w = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \cdots x_n^{\varepsilon_n},$$

where $\varepsilon_i = \pm 1$ (we agree that x^1 may denote x). Observe that this spelling of a reduced word is unique, for this is the definition of equality in the set of sequences S.

The idea of the construction of the free group F is just this: The elements of F are the reduced words and the binary operation is juxtaposition. Unfortunately, the juxtaposition of two reduced words need not be reduced. This tiny fact incurs tedious case analyses in verifying associativity, so we use a device to bypass the tedium.

Let W be the set of reduced words in X. For each $x \in X$, consider the two functions mapping W into itself, $|x|$ and $|x^{-1}|$, defined as follows:

$$|x^{\varepsilon}|(x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n}) = \begin{cases} x^{\varepsilon} x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} & \text{if } x^{\varepsilon} \neq x_1^{-\varepsilon_1}; \\ x_2^{\varepsilon_2} \cdots x_n^{\varepsilon_n} & \text{if } x^{\varepsilon} = x_1^{-\varepsilon_1}, \end{cases}$$

where $\varepsilon = \pm 1$.

Since $|x|\,|x^{-1}|$ and $|x^{-1}|\,|x|$ are each equal to the identity on W, each $|x|$ is a permutation of W (with inverse $|x^{-1}|$). Let S_W be the group of all permutations of W, and let F_0 be the subgroup of S_W generated by $X_0 = \{|x| : x \in X\}$. We claim that F_0 is free on X_0.

An arbitrary element (other than the identity) of F_0 has a factorization $|x_1^{\varepsilon_1}| \cdots |x_n^{\varepsilon_n}|$, where $\varepsilon_i = \pm 1$ and $|x^{\varepsilon}|$ and $|x^{-\varepsilon}|$ are never adjacent (otherwise we may cancel). Such a factorization is unique, for applying this function to the reduced word 1 yields the reduced word $x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n}$, and we have already noted that a reduced word has a unique spelling. Hence, two different factorizations are distinct permutations.

Suppose G is a group and $f: X_0 \to G$ is a function. Define $h: F_0 \to G$ by

$$h(|x_1^{\varepsilon_1}| \cdots |x_n^{\varepsilon_n}|) = f(|x_1|)^{\varepsilon_1} \cdots f(|x_n|)^{\varepsilon_n};$$

h is well defined, since the factorization of an element of F_0 into powers of $|x|$ is unique. Now if α and β are in F_0, then

$$\alpha = |x_1^{\varepsilon_1}| \cdots |x_n^{\varepsilon_n}| \qquad \text{and} \qquad \beta = |y_1^{\eta_1}| \cdots |y_m^{\eta_m}|;$$

thus, after possible cancellations,

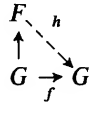
$$\alpha\beta = |x_1^{\varepsilon_1}| \cdots |x_s^{\varepsilon_s}|\,|y_t^{\eta_t}| \cdots |y_m^{\eta_m}| \qquad s \leq n,\ 1 \leq t$$

(our notation is careless, for we do not wish to overlook the cases in which all the $|x|$ are cancelled by $|y|$ or all the $|y|$ are cancelled by $|x|$). Since cancellation in F_0 implies cancellation in the group G, we have $h(\alpha\beta) = h(\alpha)h(\beta)$. Thus, h is a homomorphism; it is uniquely determined by f because X_0 generates F_0. Therefore, F_0 is free on X_0.

By Exercise 1.19, the set of reduced words in X under juxtaposition is a group isomorphic to F_0, and it is free on X. ∎

Corollary 11.2 *Every group G is a quotient of a free group.*

Proof Consider G as a set, and let F be free on G.



If $f\colon G \to G$ is the identity map, then there is a homomorphism $h\colon F \to G$ extending f; h is onto because f is onto. ∎

Definition A group G is defined by **generators** $X = \{x_k : k \in K\}$ and **relations** $\Delta = \{r_j = 1 : j \in J\}$ in case F is free on X, R is the *normal* subgroup of F generated by $\{r_j : j \in J\}$, and $G \cong F/R$. We say that $(X \mid \Delta)$ is a **presentation** of G.

Note that we are forced to use the normal subgroup generated by $\{r_j : j \in J\}$ because we want to form the quotient group F/R.

Example 5 $G = \sigma(6)$ has generators x and y and relations $x^2 = 1$, $y^3 = 1$, and $x^{-1}y^{-1}xy = 1$.

Example 6 The dihedral group D_n has generators s and t and relations $s^n = 1$, $t^2 = 1$, and $tst = s^{-1}$ (compare D_3 with $\sigma(6)$).

Example 7 The quaternions have generators a and b and relations $a^4 = 1$, $b^2 = a^2$, and $b^{-1}ab = a^{-1}$.

Example 8 A free group on $\{x_k : k \in K\}$ has generators $\{x_k : k \in K\}$ and no relations.

The existence of a group described by specifying its generators and relations has now been established. The description, however, is incomplete in that the order of the constructed group may be difficult to determine. This is not a minor difficulty, for it is even an unsolvable problem, starting with a presentation of a group G, whether or not G has only one element.

EXERCISES

11.1. Prove the existence of the following groups: (a) the generalized quaternion groups Q_n of order 2^n; (b) the groups of order p^3 described in Exercise 5.42.

11.2. A free group on no generators has one element; a free group on one generator is infinite cyclic; a free group on more than one generator is a centerless group in which every element (save the identity) has infinite order.

11.3. The group G with the presentation

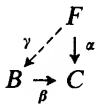
$$(x,y \mid x^2 = 1, \quad y^3 = 1)$$

is infinite.

***11.4.** If X is a set, there exists a free semigroup on X; every semigroup is a homomorphic image of a free semigroup. Conclude that an arbitrary semigroup can be described by a presentation giving generators and relations. (This existence proof is much simpler than that for groups, for there is no cancellation here. In particular, the associative law can be proved directly.) Note that since a free semigroup has no inverses, the relations in a presentation $(X \mid \Delta)$ of a semigroup are equations of the form $A_i = B_i$, where A_i and B_i are words on X having no negative exponents.

11.5. A subgroup H of G is a **retract** of G if there is a homomorphism $\varphi: G \to H$ (called a **retraction**) with $\varphi(x) = x$ for all $x \in H$. Prove that $\ker \varphi$ is the normal subgroup of G generated by $\{g\varphi(g^{-1}) : g \in G\}$ and that G is a semidirect product of $\ker \varphi$ by H (compare Exercise 9.8).

Theorem 11.3 (Projective Property) *Let $\beta: B \to C$ be onto. If F is free and $\alpha: F \to C$, then there is a map $\gamma: F \to B$ with $\beta\gamma = \alpha$.*



Proof Let $X = \{x_k: k \in K\}$ be a basis of F. Since β is onto, for each k there is an element $b_k \in B$ with $\beta(b_k) = \alpha(x_k)$. Define a function $f: X \to B$ by $f(x_k) = b_k$. There is a homomorphism $\gamma: F \to B$ with $\gamma(x_k) = b_k$ for all k since F is free on X. One checks easily that $\beta\gamma = \alpha$. ∎

Groups with the injective property are not very interesting; Baer proved that such a group must have order 1.

EXERCISES

11.6. Let F be free and K an arbitrary group; every extension of K by F is split.

****11.7.** Let F be free on $\{x_k: k \in K\}$ and let F' denote the commutator subgroup of F; F/F' is free abelian on $\{F'x_k: k \in K\}$, a set having the same number of elements as $\{x_k: k \in K\}$.

****11.8.** Let $X = \{x_k: k \in K\}$ and let $Y \subset X$. If F is free on X and H is the normal subgroup generated by Y, then F/H is free.

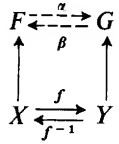
***11.9.** Show that a free group F on $\{x, y\}$ has an automorphism f with $f(f(a)) = a$ for all $a \in F$ and $f(a) = 1$ if and only if $a = 1$. (Compare Exercise 1.36.)

Now that we know the existence of free groups, let us examine their uniqueness.

Theorem 11.4 *Let X and Y be nonempty sets, F free on X, and G free on Y. Then $F \cong G$ if and only if X and Y have the same number of elements.*

Proof If $F \cong G$, then $F/F' \cong G/G'$. By Exercise 11.7, F/F' is free abelian on the cosets of X and G/G' is free abelian on the cosets of Y. By Theorem 9.17, X and Y have the same number of elements.

Suppose $f: X \to Y$ is a one-to-one correspondence. Consider the diagram



There is a homomorphism $\alpha: F \to G$ extending f and a homomorphism $\beta: G \to F$ extending f^{-1}. The composite $\alpha\beta: G \to G$ is thus an endomorphism of G fixing Y pointwise. Therefore $\alpha\beta = 1_G$, for both are homomorphisms agreeing on Y. Similarly $\beta\alpha = 1_F$, so $F \cong G$. ∎

Definition If F is free on X, then **rank** F is the number of elements in X.

Theorem 11.4 says that the definition of rank is independent of the choice of basis X.

Corollary 11.5 *If F is free on X, then F is generated by X.*

Proof Let Y be a set having the same number of elements as X and let $f: X \to Y$ be a one-to-one correspondence. Let G be the free group on Y as constructed in Theorem 11.1. In the proof of Theorem 11.4, we exhibited an isomorphism $\beta: G \to F$ with $\beta(Y) = X$. Since Y generates G, it follows that X generates F. ∎

EDGEPATH GROUPS

We are going to study an analog of Galois theory that arises in algebraic topology: there is a one-to-one correspondence between the subgroups of the fundamental group of a topological space X and certain spaces $\tilde{X}$ mapping onto X. One consequence of this study, first noticed by Baer and Levi, is that every subgroup of a free group is itself free. We mimic the topological theorems here in a completely algebraic setting.

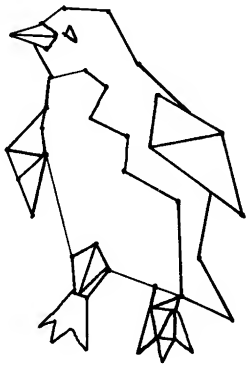
Definition A **complex** K is a set $V(K)$, called **vertices**, and a family of finite nonempty subsets of $V(K)$, called **simplexes**, such that

(i) each singleton $\{v\}$ is a simplex, where v is a vertex;
(ii) if s is a simplex, so is every nonempty subset of s.

A complex is called an *abstract simplicial complex* in the literature.

If a simplex s has $q + 1$ distinct elements, say, $s = \{v_0, v_1, \cdots, v_q\}$, we call s a **q-simplex** and say s has **dimension** q. We write dim $K = n$ and call K an **n-complex** if n is the largest dimension of a simplex in K (if there is no such largest n, we say dim $K = \infty$).

The reader should know the geometric background behind the definition of a complex, even though no geometry enters the forthcoming discussion. Think of a 0-simplex $\{v_0\}$ as a point, a 1-simplex $\{v_0, v_1\}$ as a line with endpoints v_0, v_1; a 2-simplex $\{v_0, v_1, v_2\}$ as a triangle with vertices v_0, v_1, v_2; a 3-simplex as a tetrahedron; and so forth. A complex is then a space that is built of simplexes that are assembled by gluing some of their faces together.



A complex L is a **subcomplex** of a complex K if $V(L) \subset V(K)$, and every simplex of L is also a simplex of K. For example, if s is a simplex in K, then s together with all its nonempty subsets is a subcomplex of K. Another example is provided by the **q-skeleton** of K, denoted $K^{(q)}$, defined as follows:

$$K^{(q)} = \{\text{simplexes } s \text{ in } K\text{: dim } s \leq q\}.$$

Thus, $K^{(0)} = V(K)$, $K^{(1)} = V(K) \cup \{\text{all 1-simplexes in } K\}$, and so forth. Visibly dim $K^{(q)} \leq q$. A subcomplex L of K is **full** if it contains every simplex in K all of whose vertices lie in $V(L)$. If $\{L_i\colon i \in I\}$ is a family of subcomplexes of K, then $\bigcup L_i$ and $\bigcap L_i$ are defined in the obvious way, and they are subcomplexes. In order that $\bigcap L_i$ always be defined, we allow the empty set $\varnothing$ to be a subcomplex (its dimension is -1). In particular, two subcomplexes are disjoint if and only if they have no vertices in common.

Definition An **edge** $e = (u, v)$ in K is an ordered pair of (not necessarily

distinct) vertices lying in a simplex of K; u is called the **origin** of e; v is called the **end** of e.

Definition A **path** α of **length** n is a sequence of edges

$$\alpha = e_1 \cdots e_n$$

where end e_i = orig e_{i+1}, $i = 1, \cdots, n-1$. Define **orig** α = orig e_1 and **end** α = end e_n. We say α is a path from orig α to end α. A path is **closed** at v if it is a path from v to itself.

Definition A complex K is **connected** if for every pair of vertices u, v in K there is a path in K from u to v.

EXERCISES

11.10. A subcomplex L of K is connected if and only if $L \cap K^{(1)}$ is connected.

****11.11.** Call two vertices u, v in K "connected" if there is a path in K from u to v. Prove that this defines an equivalence relation on $V(K)$.

Definition The equivalence classes defined in Exercise 11.11 are called the **components** of K.

We may consider a component C as a full subcomplex of K by regarding it as all those simplexes of K all of whose vertices lie in C. Thus, every complex K is the disjoint union of connected subcomplexes: its components.

Let $\alpha = e_1 \cdots e_n$ and $\alpha' = e'_1 \cdots e'_m$ be paths in K. If end α = orig α', then define their product

$$\alpha\alpha' = e_1 \cdots e_n e'_1 \cdots e'_m.$$

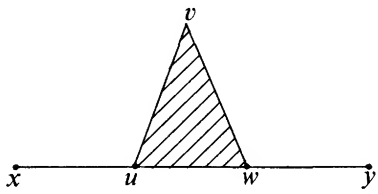
Clearly $\alpha\alpha'$, when defined, is a path from orig α to end α'. Moreover, this product is associative, when defined. We want to construct a group with the above product as multiplication. There are several obstacles: multiplication is not everywhere defined; there is no identity; there are no inverses. The last two obstacles are overcome by imposing an equivalence relation.

Definition Two paths α and α' are **equivalent**, denoted $\alpha \sim \alpha'$, if one

can be obtained from the other by a finite number of elementary moves consisting of replacing one side of an equation

$$(u, v)(v, w) = (u, w)$$

by the other whenever $\{u, v, w\}$ is a simplex of K.



For example, let K be the 2-complex pictured above, let $\alpha = (x, u)(u, w)(w, y)$ and $\alpha' = (x, u)(u, v)(v, w)(w, y)$. If K contains the shaded interior of the triangle $\{u, v, w\}$, then $\alpha \sim \alpha'$; if K does not contain the interior, then $\alpha \nsim \alpha'$.

It is easy to check that $\alpha \sim \alpha'$ defines an equivalence relation on the set of all paths in K; denote the equivalence class of α by $[\alpha]$.

EXERCISES

11.12. If $\alpha \sim \alpha'$, then orig α = orig α' and end α = end α'. Conclude that orig$[\alpha]$ and end$[\alpha]$ make sense.

11.13. If $\alpha \sim \alpha'$, $\beta \sim \beta'$, and end α = orig β, then $[\alpha\beta] = [\alpha'\beta']$.

****11.14.** Let $\alpha = \alpha_1\beta\alpha_2$, where β is a closed path lying wholly within a simplex s of K. Prove that $\alpha \sim \alpha_1\alpha_2$.

If $v \in V(K)$, set $i_v = (v, v)$; if $e = (u, v)$, set $e^{-1} = (v, u)$; if $\alpha = e_1 \cdots e_n$, set $\alpha^{-1} = e_n^{-1} \cdots e_1^{-1}$.

Lemma 11.6 *If K is a complex, let $\pi(K)$ be the set of all $[\alpha]$, where α is a path in K. Then*

(i) *each $[\alpha]$ has an origin u and end v (where u and v are vertices of K) and*

$$[i_u][\alpha] = [\alpha] = [\alpha][i_v];$$

(ii) *the associative law holds when defined;*

(iii) *if orig$[\alpha] = u$ and end$[\alpha] = v$, then*

$$[\alpha][\alpha^{-1}] = [i_u] \quad \text{and} \quad [\alpha^{-1}][\alpha] = [i_v].$$

Proof Completely without difficulty. ∎

We force $\pi(K)$ to be a group by choosing a **base point** v_*.

Definition Choose a vertex v_* in K. Then

$$\pi(K, v_*) = \{[\alpha] \in \pi(K): \alpha \text{ is a closed path at } v_*\}.$$

Theorem 11.7 *$\pi(K, v_*)$ is a group.*

Proof This follows immediately from Lemma 11.6 since multiplication is now always defined. ∎

$\pi(K, v_*)$ is called the **edgepath group** of K (it is isomorphic to the fundamental group of the topological space one may construct from K).

Corollary 11.8 *Let K be a connected complex with vertices v_* and u_*. Then*

$$\pi(K, v_*) \cong \pi(K, u_*).$$

Proof Since K is connected, there is a path β in K from v_* to u_*. Define $f: \pi(K, v_*) \to \pi(K, u_*)$ by $[\alpha] \to [\beta^{-1}\alpha\beta]$. Note that the multiplication $\beta^{-1}\alpha\beta$ takes place in $\pi(K)$, but that $[\beta^{-1}\alpha\beta] \in \pi(K, u_*)$. It is a simple matter to verify that f is a homomorphism with inverse $[\alpha'] \to [\beta\alpha'\beta^{-1}]$, using Lemma 11.6. ∎

A 1-complex K is often called a **graph**. Our next aim is to prove that $\pi(K, v_*)$ is free whenever K is a connected graph.

Definition A path $\alpha = e_1 \cdots e_n$ is **reduced** if, in this expression, no edge is adjacent to its inverse. A **circuit** is a reduced closed path.

Definition A **tree** in K is a connected subcomplex T of K with $\dim T \leq 1$ and which contains no circuits (the only tree of dimension 0 consists of a single vertex); a **maximal tree** is a tree contained in no larger tree.

It is easy to prove that an ascending union of trees is a tree, so it follows from Zorn's lemma that every complex contains a maximal tree. Note also that if u and v are distinct vertices in a tree T, then there is a unique reduced path in T from u to v (lest T contain a circuit).

Theorem 11.9 *If K is a connected complex, then a tree T in K is a maximal tree if and only if T contains all the vertices of K.*

Proof Suppose there is a vertex v not in T. Choose a vertex v_0 in T. Since K is connected, there is a path $e_1 \cdots e_n$ in K from v_0 to v.

Let $e_i = (v_{i-1}, v_i)$. We know that $v_0 \in T$ and $v_n = v \notin T$, so there must be an i with $v_{i-1} \in T$ and $v_i \notin T$. Consider the subcomplex T' with $V(T') = V(T) \cup \{v_i\}$ and with simplexes those of T together with $\{v_{i-1}, v_i\}$. Then T' is connected and contains no circuits, for any circuit in T' must contain v_i, and any path through v_i must contain $(v_{i-1}, v_i)(v_i, v_{i-1})$ or $(v_i, v_{i-1})(v_{i-1}, v_i)$. Therefore T' is a tree, contradicting the maximality of T. Thus, T contains every vertex of K.

The converse is easy and is left to the reader. ∎

EXERCISES

**11.15. Let $[\alpha] \in \pi(K, v_*)$. If each edge of α lies in a tree, then $[\alpha] = 1$.

**11.16. Let T be a finite tree. If $n_0(T)$ is the number of vertices in T and $n_1(T)$ is the number of 1-simplexes in T, then[2]

$$n_0(T) - n_1(T) = 1.$$

Given a connected complex K and a maximal tree T, define a group $G_{K,T}$ having the following generators and relations:

generators: all edges (u, v) in K;
relations: (a): $(u, v) = 1$ if (u, v) is an edge in T;
(b): $(u, v)(v, w) = (u, w)$ if u, v, w lie in a simplex of K.

Lemma 11.10 *If K is connected with maximal tree T, then*

$$\pi(K, v_*) \cong G_{K,T}.$$

REMARK Since K is connected, we know that different choices of basepoint v_* yield isomorphic groups.

Proof Let F be the free group with basis all edges (u, v) of K, and let R be the normal subgroup of relations, so that $G_{K,T} = F/R$.

For each vertex $v \neq v_*$ in K, there is a unique reduced path α_v in T from v_* to v (since T is a maximal tree); define $\alpha_{v_*} = (v_*, v_*)$.

Define a map $G_{K,T} \to \pi(K, v_*)$ as follows. First, define a function on the basis of F by

$$(u, v) \to [\alpha_u(u, v)\alpha_v^{-1}]$$

[2] If $\dim K = 1$ and K is finite, then the number $n_0(K) - n_1(K)$ is called the **Euler-Poincaré characteristic** of K.

(which is the class of a closed path at v_*). This function defines a homomorphism $\varphi\colon F \to \pi(K, v_*)$ which we claim kills the relations R. *Type* (*a*): if (u, v) is an edge in T, then $\alpha_u(u, v)\alpha_v^{-1}$ lies wholly in T, and so $\varphi((u, v)) = [\alpha_u(u, v)\alpha_v^{-1}] = 1$, by Exercise 11.15. *Type* (*b*): if $\{u, v, w\}$ is a simplex of K, then

$$\begin{aligned}[\alpha_u(u, v)\alpha_v^{-1}][\alpha_v(v, w)\alpha_w^{-1}] &= [\alpha_u(u, v)\alpha_v^{-1}\alpha_v(v, w)\alpha_w^{-1}] \\ &= [\alpha_u(u, v)(v, w)\alpha_w^{-1}] \\ &= [\alpha_u(u, w)\alpha_w^{-1}].\end{aligned}$$

Therefore φ induces a homomorphism

$$\bar{\varphi}\colon G_{K,T} \to \pi(K, v_*)$$

by

$$(u, v)R \to \varphi(u, v) = [\alpha_u(u, v)\alpha_v^{-1}].$$

We prove $\bar{\varphi}$ is an isomorphism by constructing its inverse. If $\alpha = e_1 \cdots e_n$ is a closed path in K at v_*, define

$$\theta(\alpha) = e_1 \cdots e_n R \in G_{K,T}.$$

Observe that if α' is another path in K with $\alpha \sim \alpha'$, then the relations in $G_{K,T}$ of type (b) show $\theta(\alpha) = \theta(\alpha')$. There is thus a homomorphism

$$\bar{\theta}\colon \pi(K, v_*) \to G_{K,T}$$

defined by

$$[e_1 \cdots e_n] = [\alpha] \to \theta(\alpha) = e_1 \cdots e_n R.$$

Let us compute. If $[\alpha] \in \pi(K, v_*)$ and $\alpha = e_1 \cdots e_n$, then

$$\begin{aligned}\bar{\varphi}\bar{\theta}[\alpha] &= \bar{\varphi}(\theta(\alpha)) \\ &= \bar{\varphi}(e_1 \cdots e_n R) \\ &= [\varphi(e_1) \cdots \varphi(e_n)] \text{ (since } \varphi \text{ is a homomorphism killing } R) \\ &= [\alpha_{v_*} e_1 \cdots e_n \alpha_{v_*}^{-1}].\end{aligned}$$

But $\alpha_{v_*} = (v_*, v_*)$, so that $[\alpha_{v_*}] = 1$ in $\pi(K, v_*)$. Therefore, $\bar{\varphi}\bar{\theta}$ is the identity. Finally, suppose (u, v) is a generator of $G_{K,T}$. Then

$$\begin{aligned}\bar{\theta}\bar{\varphi}((u, v)R) &= \bar{\theta}(\varphi(u, v)) \\ &= \bar{\theta}[\alpha_u(u, v)\alpha_v^{-1}] \\ &= \alpha_u(u, v)\alpha_v^{-1} R.\end{aligned}$$

Now α_v^{-1} and α_u lie in R, since their edges do, so that

$$\alpha_u(u, v)\alpha_v^{-1}R = \alpha_u(u, v)R = (u, v)R,$$

the last equation because R is normal in F (Exercise 2.37). Thus $\bar{\theta}\bar{\varphi}$ fixes a set of generators of $G_{K,T}$, hence is the identity. Therefore $\pi(K, v_*) \cong G_{K,T}$. ∎

Theorem 11.11 *If K is a connected 1-complex, then $\pi(K, v_*)$ is a free group. Moreover, if T is a maximal tree in K, then*

$$\text{rank } \pi(K, v_*) = |\{1\text{−simplexes } s\colon s \in K, s \notin T\}|.$$

Proof By Lemma 11.10, we may examine $G_{K,T}$. Because of relations of type (a), $G_{K,T}$ is surely generated by all edges (u, v), where (u, v) is not in T. Furthermore, if $e = (u, v)$, the type (b) relation $ee^{-1} = 1$ allows us to choose just one of the two edges determined by the 1-simplex $\{u, v\}$ as part of a generating set of $G_{K,T}$. Next, if u, v, w all lie in a simplex of K, then two of the vertices must be the same, for dim $K = 1$. Thus, the relations of type (b) have the form:

$$(u, u)(u, v) = (u, v);$$

$$(u, v)(v, v) = (u, v);$$

$$(u, v)(v, u) = (u, u).$$

But all of these are trivial: since $(v, v) = 1$ and $(u, v) = (v, u)^{-1}$, the subgroup R of relations is $\{1\}$. Therefore, $G_{K,T}$ is free on the 1-simplexes not in T. ∎

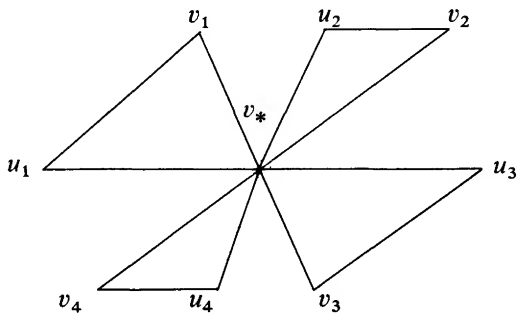
We have established a new criterion for a group to be free: if it is the edgepath group of a connected graph. Note that, in this case, a basis can be given explicitly: if T is a maximal tree in K and α_v is the unique reduced path in T from v_* to v, then a basis for $\pi(K, v_*)$ is the set of all $[\alpha_u(u, v)\alpha_v^{-1}]$, where $\{u, v\}$ is a 1-simplex not in T. Of course, for any 1-simplex $\{u, v\}$ not in T, only one edge (u, v) is chosen.

Theorem 11.12 *Given a set I, there exists a connected 1-complex K with $\pi(K, v_*)$ free of rank $|I|$.*

Proof For each $i \in I$, choose distinct points u_i, v_i, and let V be the disjoint union of a point v_* and all $\{u_i, v_i\}$. Let K be the complex with vertices V and 1-simplexes $\{v_*, u_i\}$, $\{v_*, v_i\}$, and $\{u_i, v_i\}$, all $i \in I$. Visibly K is a connected 1-complex. A maximal tree consists

of all simplexes $\{v_*, u_i\}$ and $\{v_*, v_i\}$. Therefore, $\pi(K, v_*)$ is free on all 1-simplexes $\{u_i, v_i\}$, a set in one-to-one correspondence with I. ∎

The complex K just constructed is called a **bouquet of circles** because of the picture



COVERING COMPLEXES AND THE NIELSEN-SCHREIER THEOREM

In the preceding section, we associated a group to every connected complex; in this section we associate a complex to a group.

Definition Let K and K' be complexes. A **map**[3] $f: K \to K'$ is a function $f: V(K) \to V(K')$ such that $\{fv_0, \cdots, fv_q\}$ is a simplex in K' whenever $\{v_0, \cdots, v_q\}$ is a simplex in K.

Of course, not all the vertices $fv_0, \cdots, fv_q$ need be distinct. When maps may be composed, their composite is again a map.

Definition If $p: K \to K'$ is a map and L' is a subcomplex of K', then $p^{-1}(L')$ is the subcomplex of K consisting of all those simplexes s in K with $p(s)$ in L' (in other words, $p^{-1}(L')$ is the full subcomplex of K on $p^{-1}V(L')$).

EXERCISES

11.17. If $f: K \to K'$ is a map, then the image of f is a subcomplex of K'; if K is connected, then im f is connected.

[3] Maps are called **simplicial maps** in the literature.

11.18. Let K be a complex and let $\{X_i : i \in I\}$ be a partition of $V(K)$. Define a complex $\bar{K}$ (the **quotient complex) by

$$V(\bar{K}) = \{X_i : i \in I\};$$

$\{X_{i_0}, \cdots, X_{i_q}\}$ is a simplex if there exist vertices $v_j \in X_{i_j}$ with $\{v_0, \cdots, v_q\}$ a simplex in K.

Prove that $\bar{K}$ is a complex and the function $f: V(K) \to V(\bar{K})$ sending each $v \in V(K)$ into the unique X_i containing it is a map. Moreover, a subcomplex $\bar{L}$ of $\bar{K}$ is connected if and only if $f^{-1}(\bar{L})$ is connected.

**11.19. Let K be connected and let L be a subcomplex that is a disjoint union of trees. Prove that L is contained in a maximal tree of K. (HINT: Use Exercise 11.18.)

11.20. Let I_n be the 1-complex with vertices $\{t_0, t_1, \cdots, t_n\}$ and 1-simplexes $\{t_0, t_1\}, \{t_1, t_2\}, \cdots, \{t_{n-1}, t_n\}$. Prove that a path in K from u to v is a map $\alpha: I_n \to K$ (some n) with $\alpha(t_0) = u$ and $\alpha(t_n) = v$.

11.21. Let $f: K \to K'$ be a map with $f(v_*) = v'_*$. Prove that

$$f_\#: \pi(K, v_*) \to \pi(K', v'_*)$$

defined by

$$[\alpha] \to [f \circ \alpha]$$

is a homomorphism, where $\alpha: I_n \to K$ is a path with $\alpha(t_0) = v_* = \alpha(t_n)$. Note that if $\alpha = (v_*, v_1)(v_1, v_2) \cdots (v_n, v_*)$, then $f \circ \alpha = (v'_*, fv_1)(fv_1, fv_2) \cdots (fv_n, v'_*)$.

11.22. If $f: K \to K$ is the identity, then $f_\#$ is the identity. Given maps $f: K \to K'$ and $g: K' \to K''$ with $f(v_*) = v'_*$ and $g(v'_*) = v''_*$, then $(g \circ f)_\# = g_\# \circ f_\#: \pi(K, v_*) \to \pi(K'', v''_*)$.

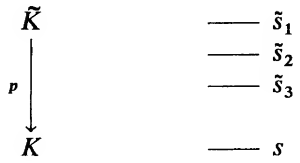
Definition Let K be a complex. A pair $(\tilde{K}, p)$ is a **covering complex** of K if

(i) $p: \tilde{K} \to K$ is a map;
(ii) $\tilde{K}$ is connected;
(iii) for every simplex s in K, $p^{-1}(s)$ is a union of pairwise disjoint simplexes,

$$p^{-1}(s) = \bigcup \tilde{s}_i,$$

with $p|\tilde{s}_i: \tilde{s}_i \to s$ a one-to-one correspondence for each i.

The map $p\colon \tilde{K} \to K$ is called the **projection** and the simplexes $\tilde{s}_i$ are called the **sheets** over s. The picture to keep in mind is



Example 9 Let K be a "triangle": a 1-complex with vertices $\{v_0, v_1, v_2\}$ and 1-simplexes $\{v_0, v_1\}$, $\{v_1, v_2\}$ and $\{v_0, v_2\}$. Let $\tilde{K}$ be the 1-complex with vertices $\{t_i : i \in Z\}$ and 1-simplexes $\{t_i, t_{i+1}\}$, $i \in Z$. If we define $p\colon \tilde{K} \to K$ by $p(t_i) = v_j$ if $i \equiv j \pmod 3$, then $(\tilde{K}, p)$ is a covering complex of K.

EXERCISES

11.23. If $p\colon \tilde{K} \to K$ is a covering complex, then image $f = K$; conclude that if K has a covering complex, then K is connected.

****11.24.** Let $p\colon \tilde{K} \to K$ be a covering complex and let L be a connected subcomplex of K. If $\tilde{L}$ is a component of $p^{-1}(L)$, then $p|\tilde{L}\colon \tilde{L} \to L$ is a covering complex.

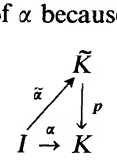
****11.25.** If $p\colon \tilde{K} \to K$ is a covering complex and T is a tree in K, then $p^{-1}(T)$ is a disjoint union of trees. (HINT: Show each component of $p^{-1}(T)$ is a tree.)

****11.26.** If $p\colon \tilde{K} \to K$ is a covering complex and s is a q-simplex in K, then every sheet $\tilde{s}_i$ over s is a q-simplex. Conclude that

$$\dim \tilde{K} = \dim K.$$

Lemma 11.13 *Let $p\colon \tilde{K} \to K$ be a covering complex and let $p(\tilde{v}_*) = v_*$. Given a path α in K with origin v_*, there exists a unique path $\tilde{\alpha}$ in $\tilde{K}$* with origin $\tilde{v}_*$ and $p\tilde{\alpha} = \alpha$.

REMARK One calls $\tilde{\alpha}$ a **lifting** of α because of the picture



Proof We prove by induction that if α has length n and origin v_*, then there exists a unique $\tilde{\alpha}$ of length n with origin $\tilde{v}_*$ and $p\tilde{\alpha} = \alpha$.

If $n = 1$, then $\alpha = (v_*, v)$, where $\{v_*, v\}$ lie in a simplex s of K. We may assume $v_* \neq v$, so that s is a 1-simplex. If $\tilde{s}_i$ is the sheet over s containing $\tilde{v}_*$, then $\tilde{s}_i$ is a 1-simplex, by Exercise 11.26. Hence $\tilde{s}_i = \{\tilde{v}_*, \tilde{v}\}$ for some vertex $\tilde{v}$ in $\tilde{K}$, so $(\tilde{v}_*, \tilde{v})$ is an edge in $\tilde{K}$ with origin $\tilde{v}_*$ and with $p\tilde{\alpha} = \alpha$. To see that $\tilde{\alpha}$ is unique, suppose $(\tilde{v}_*, \tilde{v})$ and $(\tilde{v}_*, \tilde{u})$ are liftings of α. If $s = \{v_*, v\}$, then the vertices $\tilde{v}_*$, $\tilde{v}$, and $\tilde{u}$ lie in $p^{-1}(s) = \bigcup \tilde{s}_i$. Since $p^{-1}(s)$ is the full subcomplex of K on $p^{-1}(\{v_*, v\})$, both edges $(\tilde{v}_*, \tilde{v})$ and $(\tilde{v}_*, \tilde{u})$ lie in $p^{-1}(s)$. Visibly, the sheet containing $\tilde{v}$ and the sheet containing $\tilde{u}$ are not disjoint. It follows that $\tilde{u} = \tilde{v}$.

Assume $n > 1$, so that $\alpha = \alpha_1(v, v')$, where α_1 is a path from v_* to v of length $n - 1$. By induction, there is a unique lifting $\tilde{\alpha}_1$ from $\tilde{v}_*$ to $\tilde{v}$, where $\tilde{v}$ is some vertex in $\tilde{K}$ with $p(\tilde{v}) = v$. The first step of the induction provides a unique vertex $\tilde{v}'$ in $\tilde{K}$ with $p(\tilde{v}') = v'$ and with $\{\tilde{v}, \tilde{v}'\}$ a simplex in $\tilde{K}$. Therefore $\tilde{\alpha}_1(\tilde{v}, \tilde{v}')$ is the unique lifting of α having origin $\tilde{v}_*$. ∎

Lemma 11.14 *Let $p: \tilde{K} \to K$ be a covering complex and let $p(\tilde{v}_*) = v_*$. If α and β are equivalent paths in K with origin v_*, then their liftings $\tilde{\alpha}$ and $\tilde{\beta}$ having origin $\tilde{v}_*$ are also equivalent.*

Proof Recall that equivalence means that α can be transformed into β by a finite number of elementary moves that replace one side of an equation

$$(u, v)(v, w) = (u, w)$$

by the other whenever u, v, w all lie in a simplex of K. Thus, it suffices to prove that if $(\tilde{u}, \tilde{v})(\tilde{v}, \tilde{w})$ is a lifting of $(u, v)(v, w)$ and if u, v, w lie in a simplex s of K, then $\tilde{u}$, $\tilde{v}$, $\tilde{w}$ lie in a simplex of $\tilde{K}$. Let $\tilde{s}$ be the sheet over s containing $\tilde{v}$ and let $\tilde{s}_1$ be the sheet over s containing $\tilde{u}$. Thus, $\tilde{s}_1 = \{\tilde{u}, \tilde{v}_1, \tilde{w}_1\}$, where $p\tilde{v}_1 = v$ and $p\tilde{w}_1 = w$. Now $(\tilde{u}, \tilde{v}_1)$ and $(\tilde{u}, \tilde{v})$ are both liftings of the path (u, v) having origin $\tilde{u}$. By the uniqueness assertion in Lemma 11.13, $\tilde{v} = \tilde{v}_1$. Hence $\tilde{s}$ and $\tilde{s}_1$ are not disjoint, and so $\tilde{u}$ lies in $\tilde{s}$. Similarly, $\tilde{w}$ lies in $\tilde{s}$, so that $\{\tilde{u}, \tilde{v}, \tilde{w}\} = \tilde{s}$ is a simplex in $\tilde{K}$. ∎

Theorem 11.15 *Let $p: \tilde{K} \to K$ be a covering complex with $p(\tilde{v}_*) = v_*$. Then $p_\#: \pi(\tilde{K}, \tilde{v}_*) \to \pi(K, v_*)$ is one-to-one.*

Proof Assume $[\tilde{\alpha}]$ and $[\tilde{\beta}] \in \pi(\tilde{K}, \tilde{v}_*)$ are such that $p_\#[\tilde{\alpha}] = p_\#[\tilde{\beta}]$. Then $[p\tilde{\alpha}] = [p\tilde{\beta}]$ and $p\tilde{\alpha} \sim p\tilde{\beta}$. Visibly $\tilde{\alpha}$ and $\tilde{\beta}$ are the (unique) liftings of $p\tilde{\alpha}$ and $p\tilde{\beta}$ having origin $\tilde{v}_*$. By Lemma 11.14, $\tilde{\alpha} \sim \tilde{\beta}$, i.e., $[\tilde{\alpha}] = [\tilde{\beta}]$. Therefore $p_\#$ is one-to-one. ∎

What happens to $p_{\#}\pi(\tilde{K}, \tilde{v})$ as the basepoint $\tilde{v}$ is changed?

Theorem 11.16 *Let $p\colon \tilde{K} \to K$ be a covering complex with $p(\tilde{v}) = v_*$. If $p(\tilde{u}) = v_*$, then $p_{\#}\pi(\tilde{K}, \tilde{v})$ and $p_{\#}\pi(\tilde{K}, \tilde{u})$ are conjugate subgroups of $\pi(K, v_*)$. Conversely, if H is conjugate to $p_{\#}\pi(\tilde{K}, \tilde{v})$, then $H = p_{\#}\pi(\tilde{K}, \tilde{u})$ for some $\tilde{u}$ with $p(\tilde{u}) = v_*$.*

Proof If $\tilde{\beta}$ is a path in $\tilde{K}$ from $\tilde{u}$ to $\tilde{v}$, and if $\beta = p\tilde{\beta}$, then $[\beta] \in \pi(K, v_*)$ and

$$[\beta]p_{\#}\pi(\tilde{K}, \tilde{v})[\beta]^{-1} = p_{\#}\pi(\tilde{K}, \tilde{u}).$$

Conversely, assume $H = [\alpha]p_{\#}\pi(\tilde{K}, \tilde{v})[\alpha]^{-1}$. Let $\tilde{\beta}$ be the lifting of α^{-1} with origin $\tilde{v}$. If end $\tilde{\beta} = \tilde{u}$, then $p(\tilde{u}) = v_*$. Moreover,

$$\pi(\tilde{K}, \tilde{u}) = [\tilde{\beta}^{-1}]\pi(\tilde{K}, \tilde{v})[\tilde{\beta}],$$

so $p_{\#}\pi(\tilde{K}, \tilde{u}) = [\alpha]p_{\#}\pi(\tilde{K}, \tilde{v})[\alpha^{-1}] = H$. ∎

Theorem 11.17 *Let $p\colon \tilde{K} \to K$ be a covering complex. If v_* is a vertex in K and $\tilde{v}_* \in p^{-1}(v_*)$, then the index*

$$[\pi(K, v_*)\colon p_{\#}\pi(\tilde{K}, \tilde{v})] = |p^{-1}(v_*)|.$$

Proof Denote $p_{\#}\pi(\tilde{K}, \tilde{v}_*)$ by $\tilde{\pi}$. Define a function θ from the family of all right cosets of $\tilde{\pi}$ to $p^{-1}(v_*)$ by

$$\tilde{\pi}[\alpha] \to \text{end } \tilde{\alpha},$$

where $\tilde{\alpha}$ is the lifting of α having origin $\tilde{v}_*$. We claim that θ is a one-to-one correspondence.

(i) θ is well defined.

Every representative of $\tilde{\pi}[\alpha]$ has the form

$$p_{\#}([\tilde{\beta}])[\alpha] = [p\tilde{\beta} \circ \alpha]$$

for some closed path $\tilde{\beta}$ at $\tilde{v}_*$. We want the end of the lifting of $p\tilde{\beta} \circ \alpha$ having origin $\tilde{v}_*$. But such a lifting is $\tilde{\beta}\tilde{\alpha}$, and end $\tilde{\beta}\tilde{\alpha} =$ end $\tilde{\alpha}$.

(ii) θ is onto.

Let $\tilde{u} \in p^{-1}(v_*)$ and let $\tilde{\beta}$ be a path in $\tilde{K}$ from $\tilde{v}_*$ to $\tilde{u}$ (we are using the connectedness of $\tilde{K}$). Then $\alpha = p\tilde{\beta}$ is a closed path at v_*, hence $[\alpha] \in \pi(K, v_*)$. Visibly, $\tilde{\alpha} = \tilde{\beta}$, so that

$$\theta(\tilde{\pi}[\alpha]) = \text{end } \tilde{\alpha} = \text{end } \tilde{\beta} = \tilde{u}.$$

(iii) θ is one-to-one.

If $\theta(\tilde{\pi}[\alpha]) = \theta(\tilde{\pi}[\beta])$, then end $\tilde{\alpha}$ = end $\tilde{\beta}$. Hence $\tilde{\alpha}\tilde{\beta}^{-1}$ is defined and is a closed path in $\tilde{K}$ at $\tilde{v}_*$: $[\tilde{\alpha}\tilde{\beta}^{-1}] \in \pi(\tilde{K}, \tilde{v}_*)$. Therefore

$$[\alpha][\beta]^{-1} = [p\tilde{\alpha}][p\tilde{\beta}]^{-1} = p_{\#}[\tilde{\alpha}\tilde{\beta}^{-1}] \in \tilde{\pi},$$

so that $\tilde{\pi}[\alpha] = \tilde{\pi}[\beta]$. ∎

Corollary 11.18 *Let $p: \tilde{K} \to K$ be a covering complex. If v_1 and v_2 are vertices in K, then*

$$|p^{-1}(v_1)| = |p^{-1}(v_2)|.$$

Proof Let $p(\tilde{v}_1) = v_1$ and $p(\tilde{v}_2) = v_2$; let $\tilde{\beta}$ be a path in $\tilde{K}$ from $\tilde{v}_1$ to $\tilde{v}_2$ and let $\beta = p\tilde{\beta}$. It is a simple matter to check that the following diagram commutes:

$$\begin{array}{ccc} \pi(\tilde{K}, \tilde{v}_1) & \xrightarrow{G} & \pi(\tilde{K}, \tilde{v}_2) \\ \downarrow p_{\#} & & \downarrow p_{\#} \\ \pi(K, v_1) & \xrightarrow{g} & \pi(K, v_2), \end{array}$$

where $G[\tilde{\alpha}] = [\tilde{\beta}^{-1}\tilde{\alpha}\tilde{\beta}]$ and $g[\alpha] = [\beta^{-1}\alpha\beta]$. Of course, G and g are isomorphisms, which implies that the index on the left equals the index on the right. By Theorem 11.17, we have $|p^{-1}(v_1)| = |p^{-1}(v_2)|$. ∎

EXERCISE

**11.27. Let $p: \tilde{K} \to K$ be a covering complex and suppose there are j points in $p^{-1}(v)$ for each vertex v of K. Prove there are exactly j sheets over every simplex of K.

Let us now consider existence of covering complexes. Suppose first that we have a covering complex $p: \tilde{K} \to K$ with $p(\tilde{v}_*) = v_*$. Each vertex $\tilde{v}$ of $\tilde{K}$ can be described by a path α in K having origin v_*: choose a path $\tilde{\alpha}$ in $\tilde{K}$ from $\tilde{v}_*$ to $\tilde{v}$ and define $\alpha = p\tilde{\alpha}$. Had we chosen a second path, say, $\tilde{\beta}$, from $\tilde{v}_*$ to $\tilde{v}$, then the path $\beta = p\tilde{\beta}$ is also a path in K from v_* to v; moreover, $[\alpha\beta^{-1}] = p_{\#}[\tilde{\alpha}\tilde{\beta}^{-1}] \in p_{\#}\pi(\tilde{K}, \tilde{v}_*)$. This discussion motivates the next definition.

Definition Let K be a complex with basepoint v_* and let π be a subgroup of $\pi(K, v_*)$. If α and β are paths in K with origin v_*, then

$$\alpha \equiv_{\pi} \beta \quad \text{if} \quad \text{end } \alpha = \text{end } \beta \quad \text{and} \quad [\alpha\beta^{-1}] \in \pi.$$

It is easy to see that $\alpha \equiv_\pi \beta$ defines an equivalence relation on the set of all paths in K having origin v_*; denote the equivalence class of such a path α by $\bar{\alpha}$, and denote the family of all $\bar{\alpha}$ by K_π. We shall make K_π into a complex.

Let s be a simplex in K and let α be a path with origin v_* and end in s. A **continuation** of α in s is a path $\alpha\alpha'$, where α' is a path lying wholly in s. Let

$$[s, \bar{\alpha}] = \{\bar{\beta} \in K_\pi : \beta \text{ is a continuation of } \alpha \text{ in } s\}.$$

Define the simplexes in K_π to be all $[s, \bar{\alpha}]$, where s is a simplex in K and $\bar{\alpha} \in K_\pi$ is such that end $\bar{\alpha}$ is in s.

Theorem 11.19 *Let K be a connected complex, v_* a vertex of K, and π a subgroup of $\pi(K, v_*)$. Then K_π is a complex and the function $V(K_\pi) \to V(K)$ given by $\bar{\alpha} \to$ end $\bar{\alpha}$ defines a map $p: K_\pi \to K$.*

Proof Straightforward. ∎

There is an obvious choice of a basepoint in K_π: let $\bar{v}_* = \bar{\alpha}$, where $\alpha = (v_*, v_*)$.

Lemma 11.20 *If K is a connected complex, then every path α in K with origin v_* can be lifted to a path A in K_π from $\bar{v}_* = \overline{(v_*, v_*)}$ to $\bar{\alpha}$.*

Proof Let $(v_*, v_1)(v_1, v_2)\cdots(v_{n-1}, v_n)$ be a path in K from v_* to $v_n =$ end α (which exists because K is connected). Define paths α_i by $\alpha_i = (v_*, v_1)(v_1, v_2)\cdots(v_{i-1}, v_i)$. Observe that if s is the simplex $\{v_i, v_{i+1}\}$, then $\bar{\alpha}_i$ and $\bar{\alpha}_{i+1}$ both lie in $[s, \bar{\alpha}_i]$; hence $(\bar{\alpha}_i, \bar{\alpha}_{i+1})$ is an edge in K_π. Therefore $A = (\bar{v}_*, \bar{\alpha}_1)(\bar{\alpha}_1, \bar{\alpha}_2)\cdots(\bar{\alpha}_{n-1}, \bar{\alpha}_n)$ is a path in K_π from $\bar{v}_*$ to $\bar{\alpha}_n = \bar{\alpha}$ which lifts α. ∎

Corollary 11.21 *If K is a connected complex, then K_π is connected.*

Proof There is a path in K_π from $\bar{v}_*$ to every vertex $\bar{\alpha}$ of K_π. ∎

Theorem 11.22 *Let K be a connected complex and let π be a subgroup of $\pi(K, v_*)$. Then $p: K_\pi \to K$ is a covering complex and $p_\#\pi(K_\pi, \bar{v}_*) = \pi$.*

Proof Let us first show that $p: K_\pi \to K$ is a covering complex; only condition (iii) of the definition remains to be checked.

We claim that $p|[s, \bar{\alpha}]: [s, \bar{\alpha}] \to s$ is a one-to-one correspondence. Suppose $\bar{\beta}$ and $\bar{\gamma} \in [s, \bar{\alpha}]$ and $p(\bar{\beta}) = p(\bar{\gamma})$. Then $\beta = \alpha\beta_1$ and $\gamma = \alpha\gamma_1$, where β_1, γ_1 lie wholly in s. Moreover, $\beta\gamma^{-1}$ is defined and $\beta\gamma^{-1} = \alpha\beta_1\gamma_1^{-1}\alpha^{-1} \sim \alpha\alpha^{-1} \sim 1$, by Exercise 11.14, so $[\beta\gamma^{-1}] = 1$ in $\pi(K, v_*)$. Since $1 \in \pi$, $\beta \equiv_\pi \gamma$ and $\bar{\beta} = \bar{\gamma}$. Hence $p|[s, \bar{\alpha}]$ is one-to-one. To see that $p|[s, \bar{\alpha}]$ is onto, let v be a vertex in s. If α' is a path in s from end α to v, then $\bar{\alpha}\bar{\alpha}' \in [s, \bar{\alpha}]$ and $p(\bar{\alpha}\bar{\alpha}') = v$.

Let s be a simplex in K and let w be a vertex in s. It is easy to check that $p^{-1}(s) = \bigcup [s, \bar{\alpha}]$, where the union ranges over all $\bar{\alpha} \in K_\pi$ with end $\bar{\alpha} = w$. To prove that $p: K_\pi \to K$ is a covering complex, it suffices to prove the sheets $[s, \bar{\alpha}]$ are pairwise disjoint. Assume $\bar{\gamma} \in [s, \bar{\alpha}] \cap [s, \bar{\beta}]$. Then $\gamma \equiv_\pi \alpha\alpha_1$ and $\gamma \equiv_\pi \beta\beta_1$, where α_1 and β_1 are paths lying wholly in s. The definition of equivalence gives end α_1 = end β_1, so that $\alpha_1\beta_1^{-1}$ is a path lying wholly in s; moreover, $\alpha_1\beta_1^{-1}$ is a closed path at w. Hence

$$1 = [\gamma\gamma^{-1}] = [\alpha\alpha_1\beta_1^{-1}\beta^{-1}] = [\alpha\beta^{-1}] \in \pi,$$

by Exercise 11.14. It follows that $\alpha \equiv_\pi \beta$, i.e., $\bar{\alpha} = \bar{\beta}$, and so $[s, \bar{\alpha}] = [s, \bar{\beta}]$.

We claim that $\pi = p_\#\pi(K_\pi, \tilde{v}_*)$. Let $[\alpha] \in \pi(K, v_*)$. Since $p: K_\pi \to K$ is a covering complex, there is a unique lifting $\tilde{\alpha}$ of α with origin $\tilde{v}_*$. But we constructed such a lifting A in Lemma 11.20; therefore A must be $\tilde{\alpha}$ and so end $\tilde{\alpha}$ = end A = $\bar{\alpha}$. The following statements are equivalent: $[\alpha] \in p_\#\pi(K_\pi, \tilde{v}_*)$; $[\alpha] = [pA]$, where $[A] \in \pi(K_\pi, \tilde{v}_*)$; end A = orig A = $\tilde{v}_*$; $\bar{\alpha} = \tilde{v}_*$; $[\alpha(v_*, v_*)^{-1}] \in \pi$; $[\alpha] \in \pi$. ∎

We have enough information to prove the Nielsen–Schreier theorem.

Theorem 11.23 (Nielsen-Schreier) *Every subgroup H of a free group F is itself free.*

Proof Let F be free of rank $|I|$; let K be the bouquet of $|I|$ circles constructed in Theorem 11.12, so that $\pi(K, v_*)$ may be identified with F. By Theorem 11.22, there is a covering complex $p: K_H \to K$ with $p_\#\pi(K_H, \tilde{v}_*) = H$. Since $p_\#$ is one-to-one, $\pi(K_H, \tilde{v}_*) \cong H$. Since dim K = 1, Exercise 11.26 gives dim K_H = 1. Therefore H is free, by Theorem 11.11. ∎

Theorem 11.24 *Let F be free of finite rank n and let H be a subgroup of F having finite index j. Then H is free of rank $jn - j + 1$.*

Proof For a finite graph K, let $n_0(K)$ be the number of vertices and $n_1(K)$ be the number of 1-simplexes. If T is a maximal tree in K, then Exercise 11.16 gives $n_1(T) = n_0(T) - 1$. Therefore, the number of 1-simplexes in $K - T$ is $n_1(K) - n_1(T) = n_1(K) - n_0(T) + 1$. Since T is a maximal tree, Theorem 11.9 gives $n_0(T) = n_0(K)$. Conclusion: If K is a finite graph, then $\pi(K, v_*)$ is free of rank $n_1(K) - n_0(K) + 1$.

Let K be a bouquet of n circles: $n_0(K) = 2n + 1$ and $n_1(K) = 3n$. Let $p: K_H \to K$ be the covering complex corresponding to H. By Theorems 11.17 and 11.18, $[F: H] = j$ is the number of vertices lying over a vertex of K: $n_0(K_H) = jn_0(K)$. Moreover, Exercise 11.27 gives $n_1(K_H) = jn_1(K)$. We compute:

$$\begin{aligned} n_1(K_H) - n_0(K_H) + 1 &= jn_1(K) - jn_0(K) + 1 \\ &= 3jn - j(2n + 1) + 1 \\ &= jn - j + 1. \end{aligned}$$

Therefore H is free of rank $jn - j + 1$. ∎

Suppose a finite group G has a presentation

$$(x_1, \cdots, x_n \mid r_1 = 1, \cdots, r_m = 1).$$

If F is the free group on $\{x_1, \cdots, x_n\}$ and R is the normal subgroup generated by $\{r_1, \cdots, r_m\}$, then the formula above gives the rank of R. It is possible that rank $R > m$, for fewer elements may be needed to generate R as a normal subgroup than to generate R as a subgroup (in the former generation, one is allowed to include conjugates whenever desired).

EXERCISES

11.28. Let G be a finite group that is not cyclic and let $G \cong F/S$, where F is free of finite rank. Prove that rank $S >$ rank F.

11.29. Let F be free of rank 2. Does F have a proper normal subgroup H that is finitely generated?

11.30. A free group of rank > 1 is not solvable.

11.31. Exhibit infinitely many free sets of generators of a free group of rank 2.

11.32. Let H be a group such that, for every group G, each extension of G by H is split. Prove that H is free (this is the converse of Exercise 11.7).

11.33. Let G be a finitely generated group containing a subgroup H of finite index. Prove that H is finitely generated.

11.34. Let G be a group having n generators and k relations, where $n > k$; then G contains an element of infinite order. (HINT: Map a free group on n generators onto a free abelian group on n generators, and observe what happens to the relations; use Exercise 9.42.)

The theory of complexes has given us a criterion for a group G to be free: Is G the edgepath group of a connected graph? If we do not use complexes, then the only way we now know to prove a given group is free is to exhibit a basis for it. Let us sketch another proof of the Nielsen-Schreier theorem in order to exhibit a basis for a subgroup H of a free group F. Choose a transversal of H in F, i.e., one representative from each right coset Ha of H in F. Denote the chosen representative of Ha by $\overline{Ha}$. For each $x \in X$, both $\overline{Hax}$ and $(\overline{Ha})x$ lie in the coset Hax; therefore

$$t_{a,x} = (\overline{Ha})x(\overline{Hax})^{-1}$$

is an element of H. It is not difficult to prove that the elements $t_{a,x}$ just defined generate H. The next step is to further restrict the transversal.

Definition Let F be free on X and let H be a subgroup of F. A **Schreier transversal** of H in F is a transversal S such that whenever $a = x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n}$ is a reduced word in S (where $x_i \in X$ and $\varepsilon_i = \pm 1$), then every initial segment $x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k}$, $k \leq n$, also is in S.

One proves Schreier transversals exist, and then shows that H is free on all $t_{a,x} \neq 1$ arising from a Schreier transversal. Let us prove these assertions using complexes; we shall also see that Schreier transversals arise from maximal trees in a covering complex.

Lemma 11.25 *Let $p\colon \tilde{K} \to K$ be a covering complex and let T be a maximal tree in K. If v_* is a vertex of K, then each component of $p^{-1}(T)$ contains a vertex in $p^{-1}(v_*)$.*

Proof Suppose, on the contrary, that there exists a component $\tilde{C}$ of $p^{-1}(T)$ that does not meet $p^{-1}(v_*)$. Clearly $p(\tilde{C})$ is not all of $V(K)$: $v_* \notin p(\tilde{C})$. Since K is connected, it follows that there exists an edge (u, v) in K with $u \in p(\tilde{C})$ and $v \notin p(\tilde{C})$. Since $u \in p(\tilde{C})$, there is a vertex $\tilde{u} \in \tilde{K}$ with $p\tilde{u} = u$. By Lemma 11.13, the path $\alpha = (u, v)$ may be lifted to a path $\tilde{\alpha} = (\tilde{u}, \tilde{v})$ in $\tilde{K}$. Visibly $\tilde{v} \in \tilde{C}$,

for we may connect $\tilde{v}$ to $\tilde{u}$. Hence $p\tilde{v} = v \in p(\tilde{C})$, a contradiction. ∎

Theorem 11.26 *Let F be free and let H be a subgroup of F. There is a basis X of F and a Schreier transversal $\{\overline{Ha} \in Ha: a \in F\}$ such that a basis for H consists of all $t_{a,x} = (\overline{Ha})x(\overline{Hax})^{-1}$ that are distinct from 1, where x varies over X.*

Proof Identify F with $\pi(K, v_*)$, where K is a bouquet of circles with $\pi(K, v_*) \cong F$. Let $p: K_H \to K$ be the covering complex corresponding to H. Choosing a maximal tree T determines a basis of F (Theorem 11.11). Let $\bar{v}_* = (\overline{v_*, v_*})$ be the "obvious" vertex lying over v_*, so that $p_\#\pi(K_H, \bar{v}_*) = H$. Finally, let $\tilde{T}$ be a maximal tree in K_H containing $p^{-1}(T)$ (that $\tilde{T}$ exists is seen by using Exercises 11.25 and 11.19).

If $\bar{v}$ is any vertex in K_H with $p(\bar{v}) = v_*$, let $\tilde{\gamma}_{\bar{v}}$ be the unique reduced path in $\tilde{T}$ from $\bar{v}_*$ to $\bar{v}$; note that $p_\#[\tilde{\gamma}_{\bar{v}}] \in \pi(K, v_*)$. Given a coset $H[\alpha]$, let $\tilde{\alpha}$ be the lifting of α having origin $\bar{v}_*$. If $\bar{v} = \text{end } \tilde{\alpha}$, then $\bar{v}$ lies over v_* and $[\tilde{\alpha}\tilde{\gamma}_{\bar{v}}^{-1}] \in \pi(K_H, \bar{v}_*)$. Applying $p_\#$ gives

$$H[\alpha] = Hp_\#[\tilde{\gamma}_{\bar{v}}],$$

so the set of $p_\#[\tilde{\gamma}_{\bar{v}}]$ is a transversal.

To see the set of all $p_\#[\tilde{\gamma}_{\bar{v}}]$ is a Schreier transversal, write $\tilde{\gamma}_{\bar{v}} = \tilde{\alpha}_1 \cdots \tilde{\alpha}_n$, where each $\tilde{\alpha}_i$ contains exactly one edge not in $p^{-1}(T)$. For each $i \geq 1$, let $\tilde{u}_i = \text{end } \tilde{\alpha}_i = \text{orig } \tilde{\alpha}_{i+1}$. Now $\tilde{u}_i$ lies in some component $\tilde{C}_i$ of $p^{-1}(T)$, and there is a vertex $\tilde{v}_i$ in $\tilde{C}_i$ lying over v_*, by Lemma 11.25. Since $\tilde{C}_i$ is a tree, there is a unique reduced path $\tilde{\beta}_i$ in $\tilde{C}_i$ from $\tilde{u}_i$ to $\tilde{v}_i$. Consider the new path

$$(\tilde{\alpha}_1\tilde{\beta}_1)(\tilde{\beta}_1^{-1}\tilde{\alpha}_2\tilde{\beta}_2)(\tilde{\beta}_2^{-1}\tilde{\alpha}_3\tilde{\beta}_3)\cdots(\tilde{\beta}_{n-1}^{-1}\tilde{\alpha}_n).$$

Observe that $\tilde{\alpha}_1\tilde{\alpha}_2\cdots\tilde{\alpha}_i\tilde{\beta}_i \sim \tilde{\gamma}_{\tilde{v}_i}$, for both are paths in $\tilde{T}$ from $\bar{v}_*$ to $\tilde{v}_i$. Further, for each $\tilde{\delta}_i = \tilde{\beta}_{i-1}^{-1}\tilde{\alpha}_i\tilde{\beta}_i$ we have $p\tilde{\delta}_i$ a closed path in K at v_* that contains only one edge not in T; thus $\tilde{\delta}_i$ determines a generator of $\pi(K, v_*)$. It follows that each initial segment $\tilde{\delta}_1\tilde{\delta}_2\cdots\tilde{\delta}_i$, $i \leq n$, determines a member of the transversal, for $p_\#[\tilde{\delta}_1\cdots\tilde{\delta}_i] = p_\#[\tilde{\gamma}_{\tilde{v}_i}]$.

Finally, a basis for $\pi(K_H, \bar{v}_*)$ consists of all $[\tilde{\gamma}_{\bar{u}}(\bar{u}, \bar{v})\tilde{\gamma}_{\bar{v}}^{-1}]$, where $(\bar{u}, \bar{v})$ is not in $\tilde{T}$. Since $p_\#$ is one-to-one, the images of these elements comprise a basis for H, and each is of the form $t_{a,x}$. ∎

This last result can be used to contrast free groups with free abelian groups.

Theorem 11.27 *Let F be free {on x, y}. Then F', the commutator subgroup of F, is free of infinite rank.*

Proof By Exercise 11.8, F/F' is a free abelian group with basis $F'x$ and $F'y$. Therefore every right coset of F' in F has a unique representative of the form $x^m y^n$. The assignment $\overline{F'a} = x^m y^n$ is thus a well-defined function whose range is clearly a Schreier transversal of F' in F (write $x^m y^n$ as $x \cdots xy \cdots y$).

If $n > 0$, then $\overline{F'y^n} = y^n$ while $\overline{F'y^n x} \neq y^n x$. Therefore $(\overline{F'y^n})x(\overline{F'y^n x})^{-1} \neq 1$, so there are infinitely many t distinct from 1. The theorem now follows from Theorem 11.26. ∎

Thus, a subgroup of a finitely generated group need not be finitely generated.

EXERCISES

****11.35.** Let F be free on $\{x, y\}$. Prove that F contains a subgroup that is free on $\{x, y^{-1}xy, \cdots, y^{-n}xy^n, \cdots\}$.

11.36. If F is free of rank > 1, then F' is free of infinite rank.

11.37. Let F be free on $\{x, y\}$. Define $\varphi: F \to S_3$ by $x \to (12)$ and $y \to (123)$. Exhibit a basis for $\ker \varphi$.

11.38. Let F be free on $\{a, b, c, d\}$. Prove that $[a, b][c, d]$ is not a commutator.

REMARK The following theorem does not seem to follow from the proofs given above: If F is a free group of finite rank n and if $\{x_1, \cdots, x_n\}$ generates F, then $\{x_1, \cdots, x_n\}$ is a basis. We refer the reader to the book of Lyndon and Schupp for a simple combinatorial proof; we refer the reader to Massey's book for a topological proof of a stronger result due to Gruško.

FREE PRODUCTS AND THE KUROŠ THEOREM

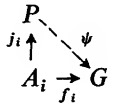
We now generalize the notion of a free group to that of a free product. As with the definition of a free group, we shall define a free product as a group having a certain "universal mapping" property. Since it is not obvious that such a group exists, we shall then be obliged to construct one. If the reader does not like the fancy definition, he is urged to look

at the existence theorem (Theorem 11.28) for a more homely, down-to-earth description of free products in terms of words.

Definition Let $\{A_i: i \in I\}$ be a family of groups. A **free product** of the A_i is a group P having the following properties:

(i) P contains an isomorphic copy of each A_i (i.e., for each i, there is a one-to-one $j_i: A_i \to P$);

(ii) for every group G and every family of homomorphisms $f_i: A_i \to G$, there is a unique homomorphism $\psi: P \to G$ extending each f_i. The diagram is



The reader should think of the maps j_i as inclusion maps and pretend that the A_i are actually subgroups of P. One should compare this definition with the analogous property of direct sums of abelian groups (Exercise 9.9).

Example 10 A free group F is a free product of infinite cyclic groups: If F is free on X, then $\langle x \rangle$ is infinite cyclic for each $x \in X$; a family of homomorphisms $f_x: \langle x \rangle \to G$ determines a function $\varphi: X \to G$ which extends to a unique homomorphism $\psi: F \to G$ with $\psi|\langle x \rangle = f_x$.

Theorem 11.28 *Given a family of groups $\{A_i: i \in I\}$, then a free product of the A_i exists.*

Proof This proof is so similar to the construction of a free group that we only present its highlights and leave the details to the reader.

Assume the groups A_i are pairwise disjoint; call $\bigcup A_i$ the alphabet and its elements letters. Form *words* with these letters; a word w is *reduced* if $w = 1$ or if $w = a_1 \cdots a_n$, where each $a_i \neq 1$ and adjacent letters lie in distinct A_i. Then a free product of the A_i has all reduced words as elements and juxtaposition as multiplication. To verify associativity, it is simplest to use the device of considering all permutations on the set of all reduced words, as in Theorem 11.1. ∎

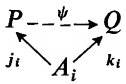
Let us describe the homomorphism ψ in the special case of the free product P of two groups A and B. Suppose G is a group and $f: A \to G$, $g: B \to G$ are homomorphisms. Then $\psi: P \to G$ is defined by

$$\psi(a_1b_1a_2b_2 \cdots a_nb_n) = f(a_1)g(b_1) \cdots f(a_n)g(b_n).$$

Having proved the existence of free products, we now prove their uniqueness.

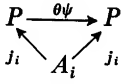
Theorem 11.29 *Let $\{A_i: i \in I\}$ be a family of groups. If P and Q are each free products of the A_i, then $P \cong Q$.*

Proof Let $j_i: A_i \to P$ and $k_i: A_i \to Q$ be the imbeddings



Since P is a free product of the A_i, there is a homomorphism $\psi: P \to Q$ with $\psi j_i = k_i$ for all $i \in I$. Similarly, there is a homomorphism $\theta: Q \to P$ with $\theta k_i = j_i$ for all $i \in I$.

Consider the new diagram



Both $\theta\psi$ and 1_P are homomorphisms $P \to P$ making the diagram commute. Since, by hypothesis, there can be only one such arrow, $\theta\psi = 1_P$. Similarly, $\psi\theta = 1_Q$ and $P \cong Q$. ∎

Because of this theorem, we may speak of *the* free product of the A_i and there is no loss in generality in regarding it as consisting of reduced words under juxtaposition, as in Theorem 11.28. The free product of $\{A_i: i \in I\}$ is denoted

$$\underset{i \in I}{*} A_i;$$

if there are only finitely many groups A_i, one usually writes $*A_i = A_1 * \cdots * A_n$.

EXERCISES

11.39. The operation of free product is commutative and associative: For any groups A, B, C, one has

$$A * B \cong B * A \quad \text{and} \quad (A * B) * C \cong A * (B * C).$$

11.40. If N is the normal subgroup of $A * B$ generated by A, then $(A * B)/N \cong B$ (cf. Exercise 11.9).

11.41. Show there is a homomorphism of $A_1 * \cdots * A_n$ onto $A_1 \times \cdots \times A_n$ which acts as the identity on each A_i.

11.42. If G' is the commutator subgroup of $G = *A_\alpha$, then $G/G' \cong \sum (A_\alpha/A'_\alpha)$ (cf. Exercise 11.8).

11.43. If G and H have more than one element, then $G * H$ is an infinite centerless group.

****11.44.** If A_α has a presentation $(X_\alpha \mid \Delta_\alpha)$, then $*A_\alpha$ has a presentation $(\bigcup X_\alpha \mid \bigcup \Delta_\alpha)$. (Assume the sets X_α, and hence the sets Δ_α, are pairwise disjoint.)

11.45. If G has a presentation $(x,y \mid x^2 = 1,\ y^3 = 1)$, then $G \cong \sigma(2) * \sigma(3)$.

11.46. Let $M = PSL(2, Z)$ be the **modular group**, i.e., the multiplicative group of all 2×2 unimodular matrices over Z modulo the subgroup $\{\pm E\}$. Prove that $M \cong \sigma(2) * \sigma(3)$ (see Exercise 2.19).

11.47. Prove that the modular group contains a free subgroup of infinite rank.

11.48. Let $A_1, \cdots, A_n, B_1, \cdots, B_m$ be indecomposable (not a direct product of proper subgroups) groups with both chain conditions. If $A_1 * \cdots * A_n \cong B_1 * \cdots * B_m$, then $n = m$ and there is a permutation σ of $\{1, \cdots, n\}$ with $A_i \cong B_{\sigma(i)}$ for all i. (The same conclusion holds if one assumes each factor is indecomposable in the sense that it is not a free product of proper subgroups.)

Kuroš has discovered the structure of the subgroups of a free product. We prove this theorem using covering complexes.

Theorem 11.30 *Let K be a connected complex with connected subcomplexes $\{K_i : i \in I\}$. Assume $\bigcup K_i = K$ and that there is a tree T in K with $K_i \cap K_j = T$, all $i \neq j$. Then*

$$\pi(K, v) \cong \underset{i \in I}{*} \pi(K_i, v_i)$$

for vertices v in K and v_i in K_i.

Proof We claim there are maximal trees T_i of K_i that contain T and such that $T' = \bigcup T_i$ is a maximal tree of K. Consider first the special case when T is a single vertex $\{t\}$. Let T_i be any maximal tree in K_i. If $T' = \bigcup T_i$ is not a tree, then it contains a shortest circuit α. Not every edge lies in one K_i lest α be a circuit in the tree T_i. It follows easily that α must pass from one K_i to another, and the only such passage is through t. Hence, α contains a shorter circuit, which is a closed path at t, contradicting the minimality of α. To verify the claim when T is a tree, let $\overline{K}$ be the quotient complex obtained from K by identifying T to a point, and let $\nu : K \to \overline{K}$ be the natural map (cf. Exercise 11.18). Our initial remarks provide

maximal trees $\overline{T}_i$ of $\overline{K}_i$ with $\overline{T}' = \bigcup \overline{T}_i$ a maximal tree of $\overline{K}$ (where $\overline{K}_i = v(K_i)$). Then $T' = v^{-1}(\overline{T}')$ is the desired maximal tree of K.

By Lemma 11.10, there is a presentation $(E_i \mid \Delta_i)$ of $\pi(K_i, v_i)$, where E_i is the set of edges in K_i and Δ_i is the set of relations of the form (a) $(u, v)(v, w) = (u, w)$ when $\{u, v, w\}$ is a simplex of K_i, and (b) $(u, v) = 1$ for every edge in T_i. By Exercise 11.44, a presentation for $*\pi(K_i, v_i)$ is $(\bigcup E_i \mid \bigcup \Delta_i)$. We claim that $(\bigcup E_i \mid \bigcup \Delta_i)$ is also a presentation of $\pi(K, v)$: $\bigcup E_i$ consists of all the edges in K; since $T' = \bigcup T_i$, an edge in K lies in T' if and only if it lies in some T_i; $\{u, v, w\}$ lies in a simplex of K if and only if it lies in a simplex of some K_i. Therefore $\pi(K, v) \cong *\pi(K_i, v_i)$. ∎

Which groups arise as $\pi(K, v)$ for some K?

Theorem 11.31 *Given a group G, there exists a connected 2-complex K^0 with $\pi(K^0, v) \cong G$.*

Proof Let G have a presentation $(X \mid \Delta)$. We construct a complex K^0 in two stages. First of all, let K be a bouquet of $|X|$ circles: specifically, K has vertices w, u_x, v_x, all $x \in X$, and 1-simplexes $\{w, u_x\}$, $\{u_x, v_x\}$, and $\{v_x, w\}$. Second, we adjoin certain 2-simplexes to K as follows: For each word $\delta \in \Delta$, we have $\delta = e_1 \cdots e_n$, where e_i is an edge in K. Let $|e_i|$ denote the underlying set of (two) vertices in e_i. The vertices of K^0 are $V(K) \cup \{z_\delta : \delta \in \Delta\}$ (one new vertex for each word $\delta \in \Delta$) and the 2-simplexes are, for each $\delta \in \Delta$, $\{z_\delta, |e_i|:$ where e_i is an edge in the factorization of $\delta\}$; the 1-simplexes in K^0 are the 1-simplexes of K and the subsets of the 2-simplexes just described. One readily sees that K^0 is a connected 2-complex and that K is a subcomplex of K^0; let $j: K \to K^0$ be the inclusion map.

The map $j: K \to K^0$ induces a homomorphism $j_\#: \pi(K, w) \to \pi(K^0, w)$. We claim that $j_\#$ is onto; to see this, it suffices to prove that every path α in K^0 closed at w is equivalent to a path involving no z_δ's. If $\alpha = e_1 \cdots e_n$, let e_i be the first edge in which a z_δ occurs. Then

$$e_i e_{i+1} = (a, z_\delta)(z_\delta, b).$$

There are two possibilities. If $a, b \in \{w, u_x, v_x\}$, then

$$(a, z_\delta)(z_\delta, b) \sim (a, b).$$

Assume $a \in \{u_x, v_x\}$ and $b \in \{u_y, v_y\}$. If $a = u_x$ and $b = v_y$, then

$$\begin{aligned}(u_x, z_\delta)(z_\delta, v_y) &\sim (u_x, z_\delta)(z_\delta, w)(w, z_\delta)(z_\delta, v_y) \\ &\sim (u_x, w)(w, v_y).\end{aligned}$$

The remaining cases are similar. Continuing in this way replaces all edges in α involving z_δ's.

Identify $\pi(K, w)$ with the free group on X, and let N be the normal subgroup generated by Δ. If we show that $N = \ker j_\#$, then $\pi(K^0, w) \cong \pi(K, w)/N \cong G$. Now $\Delta \subset \ker j_\#$, for the 2-simplexes we adjoined give relations in $\pi(K^0, w)$ making $j_\#(\delta) = 1$ for all $\delta \in \Delta$. Hence, $N \subset \ker j_\#$. To prove the reverse inclusion, let $p\colon K_N \to K$ be the covering complex corresponding to N; let $p(\tilde{w}) = w$. Suppose there are m points in K_N lying over each vertex of K (of course, m may be infinite). We construct a complex $\tilde{K}^0$ by adding to K_N m new vertices $\tilde{z}_\delta$ for each $\delta \in \Delta$

$$\begin{array}{ccc} K_N & \to & \tilde{K}^0 \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle q} \\ K & \underset{j}{\to} & K^0 \end{array}$$

and by defining 2-simplexes in the only reasonable way. This defines a covering complex $q\colon \tilde{K}^0 \to K^0$ with $K_N \subset \tilde{K}^0$. Suppose $[\alpha] \in \ker j_\#$, where $[\alpha] \in \pi(K, w)$. Now $j_\#[\alpha] = [j\alpha]$, so we are merely considering α as a path in K^0 instead of in K. Let $\tilde{\alpha}$ be the lifting of α in K_N having origin $\tilde{w}$. By hypothesis, $\alpha \sim 1$ in K^0; hence, $\tilde{\alpha} \sim 1$ in $\tilde{K}^0$, by Lemma 11.14. (Note that the lifting $\tilde{\alpha}$ of α in $\tilde{K}^0$ is the same lifting of α in K_N, by uniqueness of path lifting.) Therefore, $\tilde{\alpha}$ is a closed path at $\tilde{w}$, i.e., $[\tilde{\alpha}] \in \pi(K_N, \tilde{w})$. It follows that

$$[\alpha] = p_\#[\tilde{\alpha}] \in p_\#\pi(K_N, \tilde{w}) = N.$$

Hence $N = \ker j_\#$, as desired. ∎

Theorem 11.32 (Kuroš) *If H is a subgroup of $\underset{i \in I}{*} G_i$, then $H \cong F * (*H_\alpha)$, where F is free and each H_α is isomorphic to a subgroup of some G_i.*

Proof By Theorem 11.31, there exist connected complexes K_i with $\pi(K_i, v_i) \cong G_i$, where v_i is a vertex of K_i. Define a new complex K by adding a new vertex v_* to the disjoint union $\bigcup K_i$ and new 1-simplexes $\{v_*, v_i\}$, all $i \in I$. If T is the tree in K consisting of these new 1-simplexes, then Theorem 11.30 gives

$$\pi(K, v_*) \cong *\pi(K_i \cup T, v_i).$$

But Lemma 11.10 gives $\pi(K_i \cup T, v_i) \cong \pi(K_i, v_i) \cong G_i$. Hence

$$\pi(K, v_*) \cong *G_i.$$

Let $p: K_H \to K$ be the covering complex corresponding to H and let $\tilde{v} \in p^{-1}(v_*)$ be such that $p_\# \pi(K_H, \tilde{v}) = H$ (we have identified $*G_i$ with $\pi(K, v_*)$). For each i, $p^{-1}(K_i)$ is the disjoint union of its components $\tilde{K}_{ij}$; choose a maximal tree $\tilde{T}_{ij}$ in $\tilde{K}_{ij}$. Let $\tilde{L}$ be the 1-subcomplex of K_H:

$$\tilde{L} = \bigcup \tilde{T}_{ij} \cup p^{-1}(T).$$

Finally, let $\tilde{T}$ be a maximal tree in $\tilde{L}$ containing $\bigcup \tilde{T}_{ij}$ (which exists by Exercise 11.19). Observe that $\tilde{T}$ contains no edges in $\tilde{K}_{ij}$ aside from those in $\tilde{T}_{ij}$ lest we violate the maximality of $\tilde{T}_{ij}$ in $\tilde{K}_{ij}$.

Consider the subcomplexes $\tilde{L}$ and $\tilde{K}_{ij} \cup \tilde{T}$. Clearly K_H is the union of these, while the intersection of any two of these is the tree $\tilde{T}$. Therefore, Theorem 11.30 gives

$$\pi(K_H, \tilde{v}) \cong \pi(\tilde{L}, \tilde{v}) * (*\pi(\tilde{K}_{ij} \cup \tilde{T}, \tilde{v})).$$

Now $\pi(\tilde{L}, \tilde{v})$ is free because dim $\tilde{L} = 1$. Since $\tilde{T}$ is a maximal tree in $\tilde{K}_{ij} \cup \tilde{T}$, Lemma 11.10 gives $\pi(\tilde{K}_{ij} \cup \tilde{T}, \tilde{v}) \cong \pi(\tilde{K}_{ij}, \tilde{v}_{ij})$ for some vertex $\tilde{v}_{ij} \in \tilde{K}_{ij}$. But $p|\tilde{K}_{ij}: \tilde{K}_{ij} \to K_i$ is a covering complex, by Exercise 11.24. Hence, $\pi(\tilde{K}_{ij}, \tilde{v}_{ij})$ is isomorphic to a subgroup of $\pi(K_i, v_i) \cong G_i$. Therefore $H = p_\# \pi(K_H, \tilde{v})$ is a free product as described since $p_\#$ is one-to-one. ∎

EXERCISES

11.49. In Theorem 11.32, we showed that a subgroup H of $*G_i$ is isomorphic to $F * (*H_\alpha)$, where each H_α is a subgroup of some G_i. Prove that H is equal to $F * (*S_\alpha)$, where each S_α is a conjugate of a subgroup of some G_i (assume each G_i is imbedded in $*G_i$). (HINT: Use Theorem 11.16.)

11.50. If G is a free product of finite groups G_i, then every finite subgroup of G is isomorphic to a subgroup of some G_i.

11.51. If $G = \sigma(2) * \sigma(4)$, then $\sigma(2)$ and $\sigma(4)$ are each maximal 2-subgroups of G. Conclude that in an infinite group, sylow subgroups need not be isomorphic, let alone conjugate.

11.52. Prove that the only elements of finite order in the modular group have order 2 or 3.

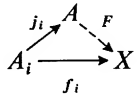
11.53. Show that the modular group contains a free subgroup of index 6. (HINT: The kernel of $\sigma(2) * \sigma(3) \to \sigma(2) \times \sigma(3)$ is torsion-free.)

11.54. Show that the commutator subgroup of the modular group is free.

****11.55.** If $f: G_1 \to G_2$ and $g: H_1 \to H_2$ are homomorphisms, there is a unique homomorphism $h: G_1 * H_1 \to G_2 * H_2$ such that $h|G_1 = f$ and $h|H_1 = g$ (compare Exercise 4.6).

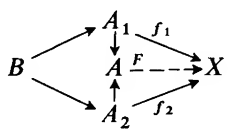
FREE PRODUCT WITH AMALGAMATED SUBGROUP

Suppose we wish to describe the union of two sets in terms of functions rather than in terms of their elements. If two sets A_1 and A_2 are disjoint, it is easy to see that their union $A = A_1 \cup A_2$ is characterized as a set containing A_1 and A_2 for which, given any set X and any pair of functions $f_i: A_i \to X$, there exists a unique function $F: A \to X$ with $F|A_i = f_i$, $i = 1, 2$. The picture is



where $j_i: A_i \to A$ is the given inclusion. We have seen this picture before. If, instead of sets and functions, we speak of abelian groups and homomorphisms, then we have described the direct sum $A_1 \oplus A_2$; if we speak of arbitrary groups and homomorphisms, then we have described the free product $A_1 * A_2$ (each of these is a special case of "coproduct" as defined in category theory).

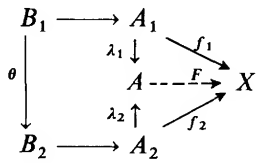
Let us now describe $A_1 \cup A_2$ when some overlap is permitted. If $B = A_1 \cap A_2$, then a moment's reflection shows that the diagrammatic characterization above remains valid if we restrict attention to those functions $f_i: A_i \to X$ that agree on B. The picture is



where the unlabeled arrows are inclusions. To say that f_1 and f_2 agree on B is to say that the diagram commutes.

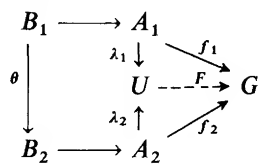
A similar diagram can be used to describe a more general set-theoretic construction. Suppose we are given two disjoint sets A_1 and A_2 with subsets $B_1 \subset A_1$ and $B_2 \subset A_2$; suppose further that there is a one-one correspondence $\theta: B_1 \to B_2$. We wish to form the union of A_1 and A_2

in which we identify B_1 and B_2 via θ (form $A_1 \cup A_2$ and identify x with $\theta(x)$ for each $x \in B_1$). The picture is



where the unlabeled arrows are inclusions. We now consider the analogous group-theoretic construction.

Definition Let A_1 and A_2 be groups containing isomorphic subgroups $B_1 \subset A_1$ and $B_2 \subset A_2$; let $\theta: B_1 \to B_2$ be an isomorphism. The **free product** of A_1 and A_2 with **amalgamated subgroup** is a group U together with homomorphisms λ_1, λ_2 such that for every group G and every pair of homomorphisms f_1, f_2 making the following diagram commute



there exists a unique homomorphism $F: U \to G$ with $F\lambda_1 = f_1$ and $F\lambda_2 = f_2$.

The reader should regard the group U as a "union" of A_1 and A_2 in which the isomorphic subgroups B_1 and B_2 are amalgamated (or identified) via the given isomorphism θ. Note that we do not assume in the definition that the homomorphisms λ_1 and λ_2 are one-one, although this will turn out to be true. We remark that one may define a free product of an arbitrary set of groups $\{A_i: i \in I\}$ with amalgamated subgroup,[4] but the special case we have defined above is the most interesting.

Theorem 11.33 *A free product with amalgamated subgroup exists.*

Proof We maintain the notation of the definition. Define

$$U = (A_1 * A_2)/N,$$

where N is the normal subgroup of $A_1 * A_2$ generated by all elements of the form $b\theta(b^{-1})$, where $b \in B_1$. Define $\lambda_i: A_i \to U$ as the com-

[4] See B. H. Neumann, *An Essay on Free Products of Groups with Amalgamations*, Phil. Trans. Royal Soc., London, 246 (1954), pp. 503–554.

posite $A_i \to A_1 * A_2 \to (A_1 * A_2)/N = U$. The easy verification that U satisfies all the requirements is left to the reader. ▮

EXERCISES

11.56. If U and V are both free products of A_1 and A_2 in which subgroups B_1 and B_2 are amalgamated via an isomorphism θ, then $U \cong V$.

***11.57.** Let A_i have a presentation $A_i = (X_i \mid \Delta_i)$, $i = 1, 2$. If Φ is the set of words on $X_1 \cup X_2$ of the form $b\theta(b^{-1})$, $b \in B_1$, then a presentation of the free product with amalgamated subgroup is

$$(X_1 \cup X_2 \mid \Delta_1 \cup \Delta_2 \cup \Phi).$$

Notation The group U that is the free product of A_1 and A_2 with subgroups amalgamated via a given isomorphism θ is denoted

$$U = A_1 *_\theta A_2.$$

It is clear that, in $A_1 *_\theta A_2$, each $b \in B_1$ is identified with $\theta(b)$; it is not clear whether other identifications are consequences of the amalgamation. For example, is it obvious whether $\lambda_i : A_i \to U$ is one-one? Is it even obvious whether $A_1 *_\theta A_2 \neq \{1\}$?

Exercise 11.57 gives an arrow-free description of $A_1 *_\theta A_2$. Our next aim is to give a more concrete description in terms of certain types of words. For each $i = 1, 2$, choose a left transversal of B_i in A_i subject only to the condition that the representative of the coset B_i is 1. Let $\bar{a}_i$ denote the chosen representative of $a_i B_i$, so that

$$a_i = \bar{a}_i b_i \qquad \text{for some } b_i \in B_i.$$

Observe that b_i is uniquely determined by a_i (assuming the elements $\bar{a}_i$ comprising the transversals have been chosen).

Definition A **normal form** is an element of $A_1 * A_2$ of the form

$$\bar{a}_1 \bar{a}_2 \cdots \bar{a}_n b,$$

where $b \in B_1$, $n \geq 0$, and adjacent $\bar{a}$ lie in distinct A_i.

In the special case that B_1 (and hence B_2) is trivial, every reduced word in the free product is a normal form.

Theorem 11.34 (Normal Form) *Let B_i be a subgroup of A_i, $i = 1, 2$, and let $\theta : B_1 \to B_2$ be an isomorphism. Every element of $U = A_1 *_\theta A_2$ has a unique expression as a normal form.*

REMARK Just as one may think of a free product as being comprised of reduced words, we may thus think of a free product with amalgamated subgroup as being comprised of normal forms.

Proof Let N be the normal subgroup of $A_1 * A_2$ generated by $\{b\theta(b^{-1}): b \in B_1\}$. Let $x_1y_1 \cdots x_ny_n \in A_1 * A_2$, where $x_j \in A_1$, $y_j \in A_2$, and only x_1 and y_n are allowed to be 1. We perform an induction on n that there is a normal form determining the same element as $x_1y_1 \cdots x_ny_n$ in $A_1 * A_2/N = U$.

Assume $n = 1$. Now $x_1 = \bar{a}_1b_1$, where $b_1 \in B_1$, so that

$$x_1y_1 = \bar{a}_1(b_1y_1) = \bar{a}_1[\theta(b_1)y_1] \text{ in } U.$$

Since $\theta(b_1)y_1 \in A_2$, we have

$$\theta(b_1)y_1 = \bar{a}_2b_2, \qquad \text{where } b_2 \in B_2.$$

Conclusion: $x_1y_1 = \bar{a}_1\bar{a}_2b_2 = \bar{a}_1\bar{a}_2\theta^{-1}(b_2)$ in U, and the last element is a normal form. The inductive step is proved in the same way. (Observe that after the process ends, the last factor b lies in B_1 or B_2; if $b \in B_1$, we have a normal form; if $b \in B_2$, replace b by $\theta^{-1}(b)$ and we have a normal form.) We have thus shown that every $u \in U$ is represented by a normal form. But we have done more: Given any reduced word $x_1y_1 \cdots x_ny_n$ in $A_1 * A_2$, the inductive process above assigns a specific normal form to it (denote this normal form $F(x_1y_1 \cdots x_ny_n)$).

In order to prove the uniqueness of this normal form, we construct a homomorphism on U having distinct values on elements described by different normal forms. We use the device of Theorem 11.1. Let M be the set of all normal forms; observe that two normal forms are equal if and only if they have the same spelling. If $a \in A_i$, define a function $|a|: M \to M$ by

$$|a|(\bar{a}_1 \cdots \bar{a}_nb) = F(a\bar{a}_1 \cdots \bar{a}_nb).$$

Clearly, $|1|$ is the identity function on M. Furthermore, consideration of the several cases (depending on possible cancellations) shows that for $a, a' \in A_i$

$$|a| \circ |a'| = |aa'|.$$

Therefore $|a^{-1}| = |a|^{-1}$, so each $|a|$ is a permutation of M. If S_M is the group of all permutations of M, then $a \to |a|$ is a homomorphism $A_i \to S_M$. In particular, if $b \in B_1 \subset A_1$, then $|b|: M \to M$ is defined.

The defining property of free product allows us to assemble these homomorphisms into a homomorphism

$$\Gamma\colon A_1 * A_2 \to S_M$$

taking a normal form $\bar{a}_1 \cdots \bar{a}_n b$ into $|\bar{a}_1| \cdots |\bar{a}_n|\,|b|$. Now $b\theta(b^{-1}) \in \ker \Gamma$ for all $b \in B_1$, so that Γ induces a homomorphism

$$\Gamma'\colon U \to S_M$$

by

$$\Gamma'(\bar{a}_1 \cdots \bar{a}_n bN) = \Gamma(\bar{a}_1 \cdots \bar{a}_n b) = |\bar{a}_1| \cdots |\bar{a}_n|\,|b|.$$

Thus Γ' assigns distinct values in S_M to distinct normal forms, for

$$|\bar{a}_1| \cdots |\bar{a}_n|\,|b|(1) = \bar{a}_1 \cdots \bar{a}_n b,$$

and the spelling of normal forms is unique. Therefore, distinct normal forms describe distinct elements of U. ∎

Theorem 11.35 *Let $U = A_1 *_\theta A_2$ be a free product with amalgamated subgroup, where $\theta\colon B_1 \to B_2$ is an isomorphism.*

(i) *The homomorphisms $\lambda_i\colon A_i \to U$ are one-one.*
(ii) *If $A_i' = \lambda_i(A_i)$, then $U = \langle A_1', A_2' \rangle$ and $A_1' \cap A_2' \cong B_1$.*

Proof (i) If $a_1 \in A_1$, its normal form is $\bar{a}_1 b$ for some $b \in B_1$; moreover, $a_1 \neq 1$ implies $\bar{a}_1 b \neq 1$, by Theorem 11.34. Let $\Gamma'\colon U \to S_M$ be the homomorphism defined in Theorem 11.34. Then $\Gamma'\lambda_1(a_1) = \Gamma'(a_1 N) = \Gamma'(\bar{a}_1 bN) = |\bar{a}_1|\,|b|$. Since $\Gamma'\lambda_1$ is one-one, it follows that λ_1 is one-one. A similar argument shows that λ_2 is also one-one.

(ii) Since $U = A_1 * A_2/N$, it is clear that $U = \langle A_1', A_2' \rangle$, where $A_i' = \lambda_i(A_i)$. If $u \in A_1' \cap A_2'$, then $\lambda_1(a_1) = u = \lambda_2(a_2)$, where $a_i \in A_i$. Let a_1 have normal form $\bar{a}_1 b$ and let a_2 have normal form $\bar{a}_2 b'$. Then applying Γ' gives $|\bar{a}_1|\,|b| = |\bar{a}_2|\,|b'|$. It follows from the uniqueness of normal form that $\bar{a}_1 = 1 = \bar{a}_2$ and that $b = b'$. Conversely, it is easy to see that $bN \in A_1' \cap A_2'$ for every $b \in B_1$. Therefore, $A_1' \cap A_1' = \{bN\colon b \in B_1\} = \lambda_1(B_1) \cong B_1$. ∎

EXERCISES

11.58. **(van Kampen)** Let L_1 and L_2 be connected complexes with $L_1 \cup L_2$ and $L_1 \cap L_2$ connected. If v is a vertex in $L_1 \cap L_2$, then $\pi(L_1 \cup L_2, v)$ is a free product of $\pi(L_1, v)$ and $\pi(L_2, v)$ with amalgamated subgroup $\pi(L_1 \cap L_2, v)$.

**11.59. Let U be a free product of groups A_1 and A_2 with amalgamated subgroup $B = A_1 \cap A_2$ (make the identification $A_i = A_i'$, $i = 1, 2$). Prove that if $y_1, \cdots, y_r$ are elements of U not in B with $y_j \in A_{i_j}$ and $i_j \neq i_{j+1}$, then $z = y_1 y_2 \cdots y_r \notin B$. Conclude that $z \neq 1$.

We now apply this last construction to obtain some imbedding theorems.

Theorem 11.36 (Higman, Neumann, and Neumann)[5] *Let G be a group that contains isomorphic subgroups A and B; let $\varphi: A \to B$ be an isomorphism. There exists a group H containing G in which φ is induced by an inner automorphism, i.e., there is an element $t \in H$ with*

$$\varphi(a) = t^{-1}at \qquad \text{for all } a \in A.$$

Proof Let $\langle u \rangle$ and $\langle v \rangle$ be disjoint infinite cyclic groups; let

$$K_1 = G * \langle u \rangle,$$

$$K_2 = G * \langle v \rangle,$$

and let L_1 be the subgroup of K_1 generated by G and $u^{-1}Au$. Now

$$L_1 = G * u^{-1}Au,$$

for there can be no equation

$$g_1 u^{-1} a_1 u g_2 u^{-1} a_2 u \cdots g_n u^{-1} a_n u = 1$$

in K_1; a fortiori, there can be no such equation in L_1. Similarly, the subgroup L_2 of K_2 generated by G and $v^{-1}Bv$ is a free product:

$$L_2 = G * v^{-1}Bv.$$

By Exercise 11.55, there is an isomorphism $\theta: L_1 \to L_2$ such that $\theta|G$ is the identity and $\theta(u^{-1}au) = v^{-1}\varphi(a)v$.

Let $H = K_1 *_\theta K_2$ be the free product of K_1 and K_2 in which we amalgamate L_1 and L_2 via the isomorphism θ. By Theorem 11.35, H contains a subgroup isomorphic to L_1 (and L_1 contains G).

Furthermore, for each $a \in A$,

$$u^{-1}au = v^{-1}\varphi(a)v.$$

[5] See Exercise 12.18 for a simpler construction of such a group H.

If $t = uv^{-1}$, then $t \in H$, and for all $a \in A$

$$t^{-1}at = \varphi(a). \quad \blacksquare$$

If G is a countable group, then G is a homomorphic image of a countable free group F of infinite rank:

$$F/R \cong G,$$

where R is a normal subgroup of F. Now we know that F can be imbedded in a free group F_1 on two generators (Theorem 11.27); were R normal in the larger group F_1, then G would be imbedded in F_1/R, a group on two generators. This proof is fictitious, but the theorem is true.

Theorem 11.37 (Higman, Neumann, and Neumann) *Every countable group G can be imbedded in a group H that can be generated by two elements.*

REMARK In Exercise 3.40, we have already noted this result in the very special case when G is finite.

Proof Let $g_0, g_1, g_2, \cdots, g_n, \cdots$ be a list of all the elements of G with $g_0 = 1$. Let $H = G * F$, where F is free on $\{x, y\}$. Consider the subgroups of H

$$A = \langle x, g_1 y^{-1} xy, \cdots, g_n y^{-n} xy^n, \cdots \rangle$$

and

$$B = \langle y, x^{-1} yx, \cdots, x^{-n} yx^n, \cdots \rangle.$$

Now A and B are isomorphic: indeed, by Exercise 11.35, each is free on the displayed set of generators. Define an isomorphism $\varphi: A \to B$ by

$$\varphi(g_n y^{-n} xy^n) = x^{-n} yx^n \qquad \text{for all } n \geq 0.$$

By Theorem 11.36, there exists a group K containing H and an element t such that

$$\varphi(a) = t^{-1}at \qquad \text{for all } a \in A.$$

We claim that $L = \langle x, t \rangle \subset K$ contains G, which will complete the proof. First of all,

$$y = \varphi(x) = t^{-1}xt \in \langle x, t \rangle.$$

Next, if $n \geq 1$, then

$$t^{-1}(g_n y^{-n} x y^n)t = \varphi(g_n y^{-n} x y^n) = x^{-n} y x^n \in \langle x, t \rangle.$$

It follows that $g_n \in \langle x, t \rangle$, all $n \geq 1$. ∎

EXERCISES

11.60. Let G be a group in which every nonidentity element has infinite order. Show that G can be imbedded in a group H having exactly two conjugacy classes (compare Exercise 3.24).

11.61. Prove that there exists a group G having two generators containing an isomorphic copy of every countable abelian group. (HINT: Use Exercise 9.46.)

11.62. Prove that there exists a group G having two generators containing an isomorphic copy of every finite group.

11.63. Let G be a finitely generated group having a finite number of defining relations. Prove that G can be imbedded in a group H having two generators and only finitely many defining relations.

CHAPTER
12

The Word Problem

STATEMENT OF THE PROBLEM

Novikov and Boone have proved that there exists a group G with an unsolvable word problem; there is thus a sequence of elementary questions about G that no one machine can answer. We shall prove this remarkable result in this chapter.

Let us begin by defining the class of groups in which we shall be interested.

Definition A group is **finitely presented** in case it has a presentation with only a finite number of generators and a finite number of relations.

(B. H. Neumann first proved that there exist finitely generated groups that are not finitely presented. We shall explicitly exhibit such a group in Theorem 12.11.)

EXERCISES

12.1. Every free group of finite rank is finitely presented.

12.2. Every finite group G is finitely presented.

12.3. Every finitely generated abelian group is finitely presented.

12.4. If G has a presentation with a finite number of relations, then G is a free product of a finitely presented group and a free group.

12.5. A group G is finitely presented if and only if there is a finite connected complex K (i.e., $V(K)$ is finite) such that $G \cong \pi(K, v_*)$ for some vertex v_* of K.

Let Q be a set of questions; a decision process for Q is a "uniform set of directions which, when applied to any of the questions in Q, produces the correct answer after a finite number of steps, never at any stage of the process leaving the user in doubt as to what to do next."[1]

Suppose G is a group having the presentation

$$G = (x_1, \cdots, x_n | r_1 = 1, \cdots, r_m = 1).$$

By means of this presentation, every (not necessarily reduced) word in the x represents an element of G. We say that the **word problem** for G is **solvable** in case there exists a decision process for the set of all questions of the form: Does the word w in the x represent the identity element of G?[2]

If a word w is written $w = x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n}$, where $\varepsilon_i = \pm 1$, then the **length** of w is defined to be n. This is in contrast to the usual definition of length which requires that w be reduced. Thus, according to our definition, the empty word has length 0, but $x_1 x_1^{-1}$ has length 2.

In order to illustrate these ideas, we show that a free group

$$G = (x_1, \cdots, x_n \mid \varnothing)$$

has a solvable word problem. Here is a decision process:

1. If the length of w is 0 or 1, proceed to step 3. If the length of $w \geq 2$, underline the first adjacent pair of letters (if there is any) of the form $x_i x_i^{-1}$ or $x_i^{-1} x_i$; if there is no such pair, underline the final two letters. Proceed to step 2.
2. If the underlined pair has the form $x_i x_i^{-1}$ or $x_i^{-1} x_i$, erase it, and proceed to step 1; otherwise, proceed to step 3.
3. If the word is empty, write $w = 1$ and stop. If the word is not empty, write $w \neq 1$ and stop.

The reader should experiment a bit with this program until he is convinced it is a decision process.

[1] W. W. Boone, "The Word Problem," Annals of Math. (1959), pp. 207–265.

[2] As we have posed it, the word problem for G depends on a particular finite presentation of G. It is true (though we shall not need the fact here) that if the word problem is solvable for some one finite presentation of G, then it is solvable for every finite presentation of G.

EXERCISES

12.6. Sketch a proof that every finite group has a solvable word problem.

12.7. Sketch a proof that every finitely generated abelian group has a solvable word problem.

12.8. Sketch proofs that if each of G and H has a solvable word problem, then the same is true of their direct product $G \times H$ and their free product $G * H$.

The proof of the Novikov-Boone theorem can be split in half. The initial portion is really mathematical logic, and is a theorem of Post[3] that there exists a finitely presented semigroup having an unsolvable word problem. The more difficult portion consists of constructing a finitely presented group G and showing that solvability of the word problem for G implies solvability of the word problem for Post's semigroup. Nowhere in the reduction of the group problem to the semigroup problem is a technical definition of unsolvability used, so that the reader knowing only our intuitive description given above can follow this part of the proof. We do, however, include a technical definition below. There are several good reasons for doing this: The word problem can be properly stated; a proof of Post's theorem can be given; the generators and relations of Post's semigroup can be accounted for; a proof of a beautiful theorem of G. Higman characterizing those finitely generated groups that can be imbedded in a finitely presented group can be given.

TURING MACHINES AND POST'S THEOREM

Let us call a subset E of a set X "enumerable" if there is a computer that can recognize every element of E and no others. Of course, the nature of such a well-behaved subset E should not depend on accidental physical constraints affecting a real computer, e.g., whether the number of memory cells may exceed the total number of atoms in the universe. We thus define an idealized computer, called a *Turing machine* (after A. Turing, 1912–1954), which abstracts the essential features of a real computer and which enumerates only those subsets E that, on intuitive grounds, "ought" to be enumerable.

Informally, a Turing machine can be pictured as a box with a tape running through it. The tape consists of a serial collection of squares,

[3] This theorem was proved, independently, by Markov; we shall give Post's proof.

which is as long to the right and to the left as desired. The box is capable of printing a finite number of symbols $s_0, s_1, \cdots, s_M$ and of being in a finite number of states $q_0, q_1, \cdots, q_N$. At any fixed moment, the box is in a state q_i and is "scanning" a particular square of the tape that bears a single symbol s_j (we may agree that s_0 means blank). The action of the machine is determined by its initial structure and by q_i and s_j; this action consists of either stopping or going into some state q_l after obeying one of the following instructions:

1. Erase the symbol s_j and print some symbol s_k.
2. Move one square to the right and scan this square.
3. Move one square to the left and scan this square.

The machine is now ready for its next move.

The machine is started in the first place by being given a tape, which may have some nonblank symbols printed on it (one to a square), and by being set to scan some one square while in "starting" state q_1. The machine may eventually stop or it may continue working indefinitely.

We proceed to the formal definitions; after each definition is stated, we shall give an informal interpretation. Let us first choose, once for all, two infinite lists of letters:

$$s_0, s_1, s_2, \cdots \quad \text{and} \quad q_0, q_1, q_2, \cdots.$$

Definition A **quadruple** is a 4-tuple of one of the following three types:

$$q_i s_j s_k q_l;$$
$$q_i s_j R q_l;$$
$$q_i s_j L q_l.$$

Definition *A* **Turing machine** is a finite, nonempty set of quadruples no two of which have the same first two letters.

The three types of quadruples correspond to the three types of moves in our informal description given above. For example, $q_i s_j R q_l$ may be interpreted as being the instruction: "When scanning symbol s_j in state q_i, move right one square and enter state q_l".

Definition A word on a given set of letters is **positive** if it has no negative exponents. We agree that the empty word is positive.

Definition An **instantaneous description** α is a positive word on letters s_j and exactly one q_i (which is not at the right end).

As an example, the instantaneous description $s_2s_0q_1s_5s_2$ is to be interpreted, "The symbols on the tape are $s_2s_0s_5s_2$ (*with blanks everywhere else*) and the machine in state q_1 scanning s_5".

Definition Let T be a Turing machine and let α, β be instantaneous descriptions. We write $\alpha \to \beta$ if there are (possibly empty) positive words P and Q on s-letters such that one of the following conditions holds:

(i) $\left.\begin{aligned}\alpha &= Pq_is_jQ\\ \beta &= Pq_ls_kQ\end{aligned}\right\}$ where $q_is_js_kq_l \in T$.

(ii) $\left.\begin{aligned}\alpha &= Pq_is_js_kQ\\ \beta &= Ps_jq_ls_kQ\end{aligned}\right\}$ where $q_is_jRq_l \in T$.

(iii) $\left.\begin{aligned}\alpha &= Pq_is_j\\ \beta &= Ps_jq_ls_0\end{aligned}\right\}$ where $q_is_jRq_l \in T$.

(iv) $\left.\begin{aligned}\alpha &= Ps_kq_is_jQ\\ \beta &= Pq_ls_ks_jQ\end{aligned}\right\}$ where $q_is_jLq_l \in T$.

(v) $\left.\begin{aligned}\alpha &= q_is_jQ\\ \beta &= q_ls_0s_kQ\end{aligned}\right\}$ where $q_is_jLq_l \in T$.

The reader should interpret $\alpha \to \beta$ as a **basic move** of the machine. Some further explanation is now needed to interpret basic moves described by condition (iii) or (v). The tape is finite, but when the machine comes to an end of the tape, the tape is lengthened by adjoining a blank square. Since s_0 means "blank", the two rules (iii) and (v) thus correspond to the cases when the machine is scanning either the last symbol or the first symbol on the tape.

The proviso in the definition of a Turing machine that no two quadruples have the same first two symbols may be interpreted to mean that there is never ambiguity about a machine's next move. Formally, $\alpha \to \beta$ and $\alpha \to \gamma$ implies $\beta = \gamma$.

Definition A **computation** of a Turing machine is a finite sequence of instantaneous descriptions $\alpha_1, \alpha_2, \cdots, \alpha_t$, where

$$\alpha_i \to \alpha_{i+1}, \qquad i = 1, \cdots, t-1,$$

and α_t is **terminal**, i.e., there is no α with $\alpha_t \to \alpha$.

Definition If T is a Turing machine, its **alphabet** is the set $\{s_0, \cdots, s_M\}$ of all s-letters occurring in its quadruples. If w is a positive word on the

alphabet of T, we say that $T(w)$ **exists** if there is a computation of T beginning with $q_1 w$.

Informally, we regard the machine T as being in starting state q_1 and as having w printed on its tape. The running of T should be visualized as a perhaps infinite sequence of instantaneous descriptions $q_1 w \to \alpha_2 \to \alpha_3 \to \cdots$. This sequence stops if $T(w)$ exists; otherwise, T runs forever.

Given a Turing machine T, let W be the set of all positive words on its alphabet. To the machine T, we associate the following subset of W:

$$e(T) = \{w \in W\colon T(w) \text{ exists}\};$$

we say that T **enumerates** $e(T)$.

Definition Let W be the set of positive words on letters $\{s_0, \cdots, s_M\}$. A subset E of W is **r.e. (recursively enumerable)** if there is some Turing machine T that enumerates E, i.e., the alphabet of T contains $\{s_0, \cdots, s_M\}$ and $E = e(T)$.

This notion of r.e. subsets of words can be specialized to subsets of the natural numbers $\mathbf{N}$ by considering just the words of the form s_1^{n+1}, where $n \geq 0$. Thus, a subset E of $\mathbf{N}$ is an r.e. subset if there is a Turing machine T with $E = \{n \in \mathbf{N}\colon T(s_1^{n+1}) \text{ exists}\}$.

EXERCISES

12.9. Prove that there exist subsets of $\mathbf{N}$ that are not r.e. (HINT: There are only countably many Turing machines.)

12.10. Prove that the set of even natural numbers is r.e.

12.11. Give an example of a Turing machine T with s_1 in its alphabet and such that $T(s_1)$ does not exist.

12.12. If W is the set of all positive words on $\{s_0, \cdots, s_M\}$ and if E_1 and E_2 are r.e. subsets of W, then $E_1 \cup E_2$ and $E_1 \cap E_2$ are r.e. subsets of W.

Every Turing machine T defines an r.e. subset $E = e(T) \subset W$, where W is the set of all positive words on its alphabet. How do we tell whether or not a word w lies in E? Feed $q_1 w$ into T and wait (i.e., perform the basic moves specified by the quadruples of T). If $w \in E$, then T will stop eventually. However, for a given w, there is no way of telling, a priori,

whether T will stop. Certainly this is unsatisfactory for an impatient man. But more: it leads to a surprising fact (Theorem 12.1).

Definition Let W be the set of positive words on $\{s_0, \cdots, s_M\}$. A subset E of W is **recursive** if both E and its complement $\bar{E}$ are r.e. subsets.

When E is recursive, there is never an "infinite wait" to decide whether a word w is in E. If T is a Turing machine enumerating E and if $\overline{T}$ is a Turing machine enumerating $\bar{E}$, then, for each $w \in W$, either $T(w)$ or $\overline{T}(w)$ exists. Thus, it can be decided in a finite length of time whether or not a given word w lies in E.

Our intuitive notion is that there is a decision process determining whether a word w lies in E if there is a uniform set of directions producing the correct answer (yes or no) after a finite number of steps, and never at any stage leaving the user in doubt as to what to do next. We propose that the recursive sets are precisely those subsets admitting a decision process. Of course, this proposition (called **Church's thesis**) cannot be proved, for it is a question of translating an intuitive notion into precise terms. There have been other attempts to formalize this notion, avoiding Turing machines altogether. Every alternative definition that has been proposed which gives a decision process for recursive sets has been proved to give a decision process for only these sets.

If E is a subset of $\mathbf{N}$, then the usual definition of E being recursive is that both E and $\mathbf{N} - E$ are r.e. subsets of $\mathbf{N}$. According to our definition, both E and $W - E$ are r.e. subsets of W, where W is the set of positive words on some alphabet. Using Exercise 12.12, the reader may show that these definitions are equivalent.

EXERCISE

**12.13. Let E_1 and E_2 be recursive subsets of a set of positive words W. Prove that $E_1 \cup E_2$ and $E_1 \cap E_2$ are recursive.

Theorem 12.1 *There exists an r.e. subset of the natural numbers* $\mathbf{N}$ *that is not recursive.*

Proof Since Turing machines are finite sets of quadruples based on the letters $R, L, q_0, q_1, \cdots; s_0, s_1, \cdots$, there are only countably many Turing machines. We now indicate how to enumerate them. Assign integer values to the letters in the following way:

$$R \leftrightarrow 0;\ L \leftrightarrow 1;\ q_0 \leftrightarrow 2;\ q_1 \leftrightarrow 4;\ q_2 \leftrightarrow 6; \cdots;$$
$$s_0 \leftrightarrow 3;\ s_1 \leftrightarrow 5;\ s_2 \leftrightarrow 7; \cdots.$$

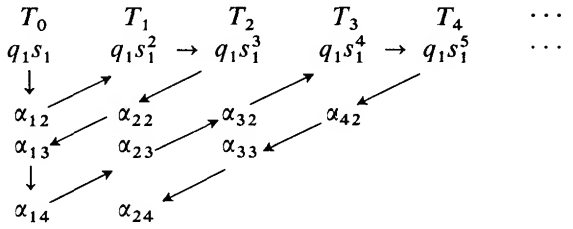
Let T be a Turing machine consisting of m quadruples. Juxtapose the quadruples in some order to form a word of length $4m$. If $p_1 < p_2 < \cdots$ is the list of all primes, define

$$G(T) = \prod_{i=1}^{4m} p_i^{e_i},$$

where e_i is the integer associated above to the ith letter in our word of length $4m$ ($G(T)$ is called the **Gödel number** of T). The fundamental theorem of arithmetic implies that distinct Turing machines have distinct Gödel numbers. If we now list these integers $G(T)$ in order of magnitude, we have effectively listed all Turing machines: $T_0, T_1, \cdots, T_n, \cdots$. Define

$$E = \{n \in \mathbf{N}: T_n(s_1^{n+1}) \text{ exists}\}.$$

We claim that E is an r.e. set. Consider the following diagram reminiscent of the proof that the set of rational numbers is countable.



The nth column consists of the sequence of basic moves of the Turing machine T_n beginning with $q_1s_1^{n+1}$. It is intuitively clear that there is an enumeration of the integers n lying in E: Follow the arrows in the diagram; if an instantaneous description α_{ni} in column n is the end of a computation, then $n \in E$. A Turing machine can, in fact, be constructed to carry out these instructions. (Moreover, for later use, there is a specific Turing machine T^* which enumerates E and which satisfies the additional condition that terminal instantaneous descriptions, and only these, involve the letter q_0.) Thus, the set E is an r.e. subset of $\mathbf{N}$.

To show that E is not recursive, it suffices to prove that

$$\bar{E} = \{n \in \mathbf{N}: n \notin E\} = \{n \in \mathbf{N}: T_n(s_1^{n+1}) \text{ does not exist}\}$$

is not an r.e. subset of $\mathbf{N}$. Suppose there were a Turing machine T enumerating $\bar{E}$. Since all Turing machines occur in the list at the

beginning of our proof, we have $T = T_{n_0}$ for some $n_0 \in \mathbf{N}$. If $n_0 \in \bar{E}$, then $n_0 \in e(T) = e(T_{n_0})$; hence, $T_{n_0}(s_1^{n_0+1})$ exists and $n_0 \in E$, a contradiction. If $n_0 \notin \bar{E}$, then $n_0 \in E$ and $T_{n_0}(s_1^{n_0+1})$ exists (definition of E); hence $n_0 \in e(T_{n_0}) = \bar{E}$, a contradiction. Therefore $\bar{E}$ is not an r.e. subset and E is not recursive. ▮

Let us link these ideas to algebra. If G is a group with generators $x_1, \cdots, x_m$, then we may regard the set W of all (not necessarily positive) words on $x_1, \cdots, x_m$ as the set of all positive words on

$$x_1, x_1^{-1}, \cdots, x_m, x_m^{-1}.$$

We say that the **word problem** for the *group* G is **solvable** if $\{w \in W: w = 1 \text{ in } G\}$ is recursive.

There is a similar definition for semigroups. If H is a semigroup with generators $x_1, \cdots, x_m$ and if W is the set of positive words on $x_1, \cdots, x_m$, then the word problem for the semigroup H is solvable if there is a decision process to determine, for an arbitrary pair of words w_1, w_2 in W whether $w_1 = w_2$ in H. This gives a criterion for unsolvability suitable for our purposes: The **word problem** for the *semigroup* H is **unsolvable** if there is a word w_0 such that $\{w \in W: w = w_0 \text{ in } H\}$ is not recursive.

Definition If U and V are (not necessarily reduced) words on some set of letters X, then

$$U \equiv V$$

means that U and V have exactly the same spelling.

Suppose a semigroup G has a presentation

$$G = (X \mid A_j = B_j, j \in J).$$

The elements of G corresponding to words U and V on X are equal if and only if there is a finite sequence of **elementary operations**

$$U \equiv W_1 \to W_2 \to \cdots \to W_t \equiv V$$

transforming U into V, where $W_i \to W_{i+1}$ consists of replacing one of the words PA_jQ and PB_jQ by the other (where P, Q are positive words on X).

We now associate a semigroup $\gamma(T)$ to a Turing machine T. Assume the letters in the quadruples of T are $q_0, \cdots, q_N, s_0, \cdots, s_M$, and let q, h

be new letters. Define $\gamma(T)$ as the semigroup having the following presentation:

generators: $q, h, q_0, q_1, \cdots, q_N, s_0, \cdots, s_M$;

relations: $q_i s_j = q_l s_k$ if $q_i s_j s_k q_l \in T$;

for all $b = 0, 1, \cdots, M$:

$$\left.\begin{aligned} q_i s_j s_b &= s_j q_l s_b \\ q_i s_j h &= s_j q_l s_0 h \end{aligned}\right\} \quad \text{if } q_i s_j R q_l \in T,$$

$$\left.\begin{aligned} s_b q_i s_j &= q_l s_b s_j \\ h q_i s_j &= h q_l s_0 s_j \end{aligned}\right\} \quad \text{if } q_i s_j L q_l \in T,$$

$$q_0 s_b = q_0,$$

$$s_b q_0 h = q_0 h,$$

$$h q_0 h = q.$$

Let us denote the set of all these relations by $R(T)$.

The first five types of relations are just the obvious ones suggested by the basic moves of the Turing machine T; the new letter h enables one to distinguish basic move (ii) from basic move (iii), and also (iv) from (v). One may thus interpret h as marking ends of the tape.

Call a word **h-special** if it is of the form $h\alpha h$ for some instantaneous description α. Since h marks ends of the tape, it is quite natural to consider h-special words.

Lemma 12.2 *Let T be a Turing machine with associated semigroup $\gamma(T)$.*

(i) *If U, V are words in $\gamma(T)$, if $U \not\equiv q$ and $V \not\equiv q$, and if $U \to V$ is an elementary operation, then U is h-special if and only if V is h-special.*

(ii) *If $U \equiv h\alpha h$, if $V \not\equiv q$, and if $U \to V$ is an elementary operation of one of the first five types, then $V \equiv h\beta h$, where either $\alpha \to \beta$ or $\beta \to \alpha$ is a basic move of T.*

Proof (i) This is true because the only relation that creates or destroys h is $hq_0h = q$.

(ii) By the first part, we know that V is h-special, say, $V \equiv h\beta h$. Now an elementary move is a substitution using an equation in a defining relation; such a relation here corresponds to a basic move of T arising from a quadruple, i.e., an elementary operation of one of

the first five types. Thus, either $\alpha \to \beta$ or $\beta \to \alpha$, since there is never ambiguity about a Turing machine's next move. ▮

Let us say that q_0 is a **stopping state** for a Turing machine T if an instantaneous description α on its alphabet is terminal if and only if α involves q_0. The Turing machine T^* described in Theorem 12.1 has q_0 as a stopping state.

EXERCISE

**12.14. Let T be a Turing machine enumerating a set E. Prove there is a Turing machine T' with the same alphabet and with stopping state q_0 that also enumerates E.

Lemma 12.3 *Let T be a Turing machine with stopping state q_0, let W be the set of positive words on the alphabet of T, and let E be the subset of W enumerated by T. For a word $w \in W$, we have*

$$w \in E \quad \textit{if and only if} \quad hq_1wh = q \quad \textit{in} \quad \gamma(T).$$

Proof If $w \in E$, there are instantaneous descriptions $\alpha_1, \cdots, \alpha_t$ with $q_1w \equiv \alpha_1 \to \alpha_2 \to \cdots \to \alpha_t$ basic moves of T and with α_t involving q_0. Using those relations in $\gamma(T)$ arising from these basic moves, one sees that $hq_1wh = h\alpha_t h$ in $\gamma(T)$. Now $\alpha_t \equiv Fq_0G$, where F and G are positive words on $s_0, s_1, \cdots, s_M$. The relations $q_0s_b = q_0$ give

$$h\alpha_t h \equiv hFq_0Gh = hFq_0h;$$

the relations $s_bq_0h = q_0h$ now give

$$hFq_0h = hq_0h = q,$$

as desired.

Assume that $hq_1wh = q$ in $\gamma(T)$. There are thus words $W_1, \cdots, W_n$ (on $\{h, q_0, q_1, \cdots, q_N, s_0, s_1, \cdots, s_M\}$) and elementary operations

$$hq_1wh \equiv W_1 \to W_2 \to \cdots \to W_n \equiv hq_0h \to q.$$

By part (i) of Lemma 12.2, each word W_i is h-special, so $W_i \equiv h\alpha_i h$ for some instantaneous description α_i. Let m be the first i for which α_i involves q_0. Note that $m \geq 2$ since α_1 involves q_1, not q_0. The proof is an induction on m. (The reason this proof differs from that of the first half is that equality in $\gamma(T)$ is, of course, a symmetric

relation, but $\alpha \to \beta$ being a basic move does not imply $\beta \to \alpha$ is a basic move.)

By part (ii) of Lemma 12.2, for each $i = 1, \cdots, m-1$, we have either $\alpha_i \to \alpha_{i+1}$ or $\alpha_{i+1} \to \alpha_i$. Also we must have $\alpha_{m-1} \to \alpha_m$ since α_m involves q_0, and q_0 is a stopping state for T. If $m = 2$, then $q_1 w \equiv \alpha_1 \to \alpha_2$ is a computation of T, so that $T(w)$ exists and $w \in E$. Assume $m > 2$. If all the arrows go to the right, i.e., if $\alpha_i \to \alpha_{i+1}$ for all i, then $q_1 w \equiv \alpha_1 \to \alpha_2 \to \cdots \to \alpha_m$ is a computation of T, and so $w \in E$. Otherwise there is an i with

$$\alpha_{i-1} \leftarrow \alpha_i \to \alpha_{i+1}.$$

Since T is a Turing machine, there is never ambiguity about its next move, and so $h\alpha_{i-1}h \equiv h\alpha_{i+1}h$. We may thus eliminate W_i, thereby reducing m. The proof is now completed by induction. ∎

Theorem 12.4 (Post) *There is a finitely presented semigroup*

$$\gamma(T^*) = (q, h, q_0, \cdots, q_N, s_0, \cdots, s_M \mid R(T^*))$$

with an unsolvable word problem. There is no decision process to determine, for an arbitrary h-special word $h\alpha h$, whether $h\alpha h = q$ in $\gamma(T^*)$.

Proof If T is a Turing machine with stopping state q_0, let A denote the alphabet of T, let W denote the set of positive words on A, and let $E \subset W$ be the subset enumerated by T. Let $\tilde{W}$ be the set of all positive words on $A \cup \{q, h, q_0, \cdots, q_N\}$, where $q_0, \cdots, q_N$ are the q-letters occurring in the quadruples of T, and let

$$\tilde{E} = \{\tilde{w} \in \tilde{W}: \tilde{w} = q \quad \text{in} \quad \gamma(T)\}.$$

Identify W with the following subset W_1 of $\tilde{W}$:

$$W_1 = \{hq_1wh: w \in W\};$$

under this identification, the subset E of W corresponds to

$$E_1 = \{hq_1wh: w \in E\}.$$

In this notation, Lemma 12.3 reads:

$$E_1 = \tilde{E} \cap W_1.$$

Assume now that $T = T^*$, a Turing machine as described in Theorem 12.1. Were $\tilde{E}$ recursive, then E_1 would also be recursive,

being an intersection of recursive sets (Exercise 12.13). This contradiction shows that $\gamma(T^*)$ has an unsolvable word problem.

To prove the final statement, let

$$\tilde{F} = \{h\text{-special words } h\alpha h\colon h\alpha h = q \quad \text{in} \quad \gamma(T^*)\}.$$

If $\tilde{F}$ is a recursive set, then so is $E_1 = \tilde{F} \cap W_1$, a contradiction. ▮

Corollary 12.5 *There is a finitely presented semigroup*

$$\Gamma = (q, q_1, \cdots, q_N, s_1, \cdots, s_M \mid F_i q_{i_1} G_i = H_i q_{i_2} K_i,\ i \in I)$$

(where F_i, G_i, H_i, K_i are (possibly empty) positive words on $S = \{s_1, \cdots, s_M\}$ and $q_{i_1}, q_{i_2} \in \{q, q_1, \cdots, q_N\}$) with an unsolvable word problem. There is no decision process to determine, for arbitrary positive words X and Y on S and for q_j, $1 \le j \le N$, whether $Xq_jY = q$ in Γ.

Proof Relabel the generators of the semigroup $\gamma(T^*)$ so that they are now $q, q_1, \cdots, q_N, s_1, \cdots, s_M$ (the q's with subscripts are reindexed so they begin with q_1 instead of q_0; h is now called s_M, and the subscripts are relabeled so that the s-letters now begin with s_1 instead of s_0). Were the set of all $Xq_jY = q$ recursive, then, in the old notation, so would the set of all h-special words $h\alpha h$ such that $h\alpha h = q$ (for it would be an intersection of recursive sets). ▮

BRITTON'S LEMMA

The history of the word problem for groups is rather involved. It was first considered by M. Dehn (1912) and A. Thue (1914). The solution was given by P. Novikov (1955) and, independently, by W. W. Boone (1954–1957). In 1959, Boone exhibited a much simpler group than any of those previously given, and he proved it has an unsolvable word problem. In 1963, J. L. Britton discovered a theorem about free products with amalgamated subgroup which shortens Boone's proof considerably. A further shortening was made in 1970 by C. F. Miller, III, and still further improvements were made by W. W. Boone and D. Collins; it is this shortest proof we present here.

In order to state Britton's result, we need the following notation.

Definition Let W be a word on $a_1, \cdots, a_n$. We say that Y is a **subword** of W if $W \equiv XYZ$ (where X or Y may be empty); a word W **involves** a_i if either a_i or a_i^{-1} is a subword of W.

If $E = (S|D)$ is a presentation of a group, we shall also denote that group by E. A word W on S determines an element of E, and that element is also denoted by W. Finally, we write

$$X = Y \quad \text{in} \quad E$$

if the words X and Y determine the same element of E.

Definition Let $E = (S \mid D)$ and $E^* = (S^* \mid D^*)$ be group presentations; we write

$$E \leq E^*$$

if $S \subset S^*$, $D \subset D^*$, and for every word W on S, we have $W = 1$ in E if and only if $W = 1$ in E^*.

EXERCISE

**12.15. Let $E = (S \mid D)$ and $E^* = (S^* \mid D^*)$, where $S \subset S^*$ and $D \subset D^*$. There is a homomorphism $\varphi: E \to E^*$ that takes any word on S into "itself"; φ is an imbedding if and only if $E \leq E^*$.

Lemma 12.6 *Let* $E = (S \mid D)$ *and*

$$E^* = (S, t \mid D, t^{-1}X_i t = X_i, \quad i \in I)$$

be presentations, where the X_i *are words on* S*; let* W *be a word on* $\{S, t\}$ *that involves* t*. If* $W = 1$ *in* E^**, then* W *contains a subword of the form* $t^e C t^{-e}$*, where* $e = \pm 1$ *and*

(i) C *is a word on* S*;*
(ii) *the group element of* E *determined by* C *lies in the subgroup of* E *generated by the* X_i*.*

Proof Let G be the subgroup of E generated by the X_i and let G' be another group isomorphic to G, say, by an isomorphism $\psi: G \to G'$. We claim that E^* is a free product with amalgamated subgroup and that $E \leq E^*$.

There is a presentation

$$G' = (x_i, \quad i \in I \mid r_j = 1, \quad j \in J),$$

where $x_i = \psi(X_i)$ and r_j is a word in the x_i. Let $\langle t \rangle$ denote an infinite cyclic group with generator t, and set $Y = G' \times \langle t \rangle$ (direct product). Finally, let A be the free product of E and Y in

which we amalgamate G and G' via the isomorphism ψ. By Exercise 11.57, one presentation of A is

$$A = (S, t, x_i, \quad i \in I \mid D, r_j = 1, \quad j \in J, \\ t^{-1}x_i t = x_i, x_i = X_i, \quad i \in I).$$

But a second presentation is

$$A = (S, t \mid D, R_j = 1, \quad j \in J, \quad t^{-1}X_i t = X_i, \quad i \in I),$$

where $\psi(R_j) = r_j$. Now each $R_j = 1$ in G and hence is 1 in E. By Exercise 12.15, we have $R_j = 1$ in A. The relations $R_j = 1$ are thus superfluous in the second presentation of A, so that $A \cong E^*$. By Theorem 11.35, the homomorphism $E \to E * Y \to E^*$ is one-to-one. Since this is the map $\varphi: E \to E^*$ of Exercise 12.15, we have $E \leq E^*$.

We now prove the lemma. If W contains a subword tt^{-1} or $t^{-1}t$, we are done. Therefore, we may assume

$$W \equiv W_0 t^{e_1} W_1 \cdots t^{e_n} W_n,$$

where $n \geq 1$, each e_j is a nonzero integer, and the W_j are words on S of which only W_0 and W_n are allowed to be empty.

If $n = 1$, then $W \equiv W_0 t^{e_1} W_1$. Since $W = 1$ in E^*,

$$t^{e_1} = W_0^{-1} W_1^{-1} \in E \cap Y = X,$$

by Theorem 11.35. This is a contradiction, for X contains no nontrivial power of t as a consequence of the amalgamation. The case $n = 1$ is thus impossible, as the lemma predicts.

For the inductive step, we use Exercise 11.59, a consequence of the normal form theorem. Applying this exercise to the word W (which is 1 in E^*), we see that W_j lies in G for some j, where $1 \leq j \leq n - 1$. If e_j and e_{j+1} have opposite sign, we are finished. Suppose e_j and e_{j+1} have the same sign. In E^*,

$$W \equiv \cdots t^{e_j} W_j t^{e_{j+1}} W_{j+1} \cdots = \cdots t^{e_j + e_{j+1}} W_j W_{j+1} \cdots.$$

The second word has one fewer occurrence of a power of t, and it satisfies all the inductive hypotheses. Therefore, it, and hence W, has a subword of the desired kind. ∎

Definition Let $E = (S \mid D)$ be a presentation. A second presentation E^* has **basis** E and **stable letters** $\{p_v : v \in V\}$ if

$$E^* = (S, p_v, \quad v \in V \mid D, p_{v_j}^{-1} A_j p_{v_j} = B_j, \quad j \in J),$$

where A_j and B_j are words on S.

For example, in Lemma 12.6, the presentation E^* has basis E and a single stable letter t.

It is possible that some $A_j = 1$ (hence $B_j = 1$). In particular, if F is free on $\{p_v : v \in V\}$, then the free product $E * F$ has basis E and stable letters $\{p_v : v \in V\}$.

For each $v \in V$, let $J(v)$ denote the following subset of the index set J:

$$J(v) = \{j \in J: p_v \text{ is involved in new relation } j\}$$
$$= \{j \in J: v = v_j \quad \text{and} \quad p_v^{-1} A_j p_v = B_j\}.$$

Given a presentation E^* with basis E, there are several subgroups of E that we must consider:

$$A = \langle A_j : j \in J \rangle \qquad B = \langle B_j : j \in J \rangle$$
$$A(v) = \langle A_j : j \in J(v) \rangle \qquad B(v) = \langle B_j : j \in J(v) \rangle.$$

For each v, $A(v)$ is a subgroup of A and $B(v)$ is a subgroup of B. We shall keep to this notation for the remainder of the chapter so that all apparently wandering A, $B(v)$, etc., really have a home.

Definition The **isomorphism condition** is that for each $v \in V$, there is an isomorphism $\varphi_v : A(v) \to B(v)$ under which $\varphi_v(A_j) = B_j$ for all $j \in J(v)$.

Observe that the isomorphism condition holds if there is an isomorphism $\varphi : A \to B$ with $\varphi(A_j) = B_j$, all $j \in J$, for then $\varphi | A(v) : A(v) \to B(v)$ is an isomorphism for all v. In particular, the isomorphism condition holds in Lemma 12.6 where φ is the identity.

Here is another example: Let E^* have the presentation

$$E^* = (w, x, y, z \mid y^{-1}xy = w, \quad y^{-1}w^{-1}xwy = xw^{-1}, \quad z^{-1}wxz = w).$$

Now E^* has basis $\langle w, x \rangle$ (which is free on these generators) and stable letters y, z. The subgroups in the definition of the isomorphism condition are:

$$A(y) = \langle x, w^{-1}xw \rangle; \qquad B(y) = \langle w, xw^{-1} \rangle;$$
$$A(z) = \langle wx \rangle; \qquad B(z) = \langle w \rangle.$$

Each of these subgroups is free on the displayed generating sets, so that the isomorphism condition holds.

EXERCISES

****12.16.** Let F be free on $\{x_1, \cdots, x_m\}$ and let $h : G \to F$ be a homomorphism. If there are elements $p_1, \cdots, p_m \in G$ with $h(p_i) = x_i$,

$i = 1, \cdots, m$, then $\langle p_1, \cdots, p_m \rangle$ is a free subgroup of G on the displayed generating set.

**12.17. Let E^* have basis E and stable letters $\{p_i : i \in I\}$. Prove that $\langle p_i : i \in I \rangle$ is a free subgroup of E^* that is free on $\{p_i : i \in I\}$. (HINT: Use Exercise 12.16.)

Lemma 12.7 *Let E^* have basis $E = (S \mid D)$ and stable letters $\{p_v : v \in V\}$. If the isomorphism condition holds, then $E \leq E^*$.*

Proof Since the isomorphism condition holds, there are isomorphisms $\varphi_v : A(v) \to B(v)$, for each v, such that $\varphi_v(A_j) = B_j$ for all $j \in J(v)$. By iterated applications of the imbedding theorem of Higman, Neumann, and Neumann (Theorem 11.36), there exists a group E_0 containing E and elements p_v, $v \in V$, such that $p_v^{-1} A_j p_v = B_j$ whenever $v \in V$ and $j \in J(v)$. If we relabel, this amounts to saying that $p_{v_j}^{-1} A_j p_{v_j} = B_j$ for all $j \in J$. Clearly, we may assume that E_0 is generated by E and p_v, $v \in V$. A presentation of E_0 can thus be obtained from that of E^* by adding (a possibly empty set of) further defining relations. Therefore, E_0 is a homomorphic image of E^*. It follows that if W is a word on S and $W = 1$ in E^*, we must have $W = 1$ in E_0, and hence $W = 1$ in E_0's subgroup E. Conversely, if $W = 1$ in E, it is trivial that $W = 1$ in E^*. We have shown that $E \leq E^*$. ∎

Definition A group E^* is a **Britton extension**[4] of E with stable letters $\{p_v : v \in V\}$ if

(i) E^* has basis E and stable letters $\{p_v, v \in V\}$ (so presentations of E^* and E are given) and
(ii) the isomorphism condition holds.

With the hypothesis of Lemma 12.6, the group E^* is a Britton extension of E. If F is a free group, then the free product $E * F$ is a Britton extension of E.

The next theorem generalizes Lemma 12.6. Note that it says something about the spelling of a (not necessarily reduced) word as well as the location of a group element corresponding to it.

Theorem 12.8 (Britton's Lemma) *Let E^* be a Britton extension of $E = (S \mid D)$ with stable letters $\{p_v : v \in V\}$. If W is a word involving*

[4] Britton extensions are often called *HNN* extensions, for they were introduced by Higman, Neumann, and Neumann.

at least one stable letter and if $W = 1$ in E^*, then W contains a subword of the form $p_v^e C p_v^{-e}$, where $e = \pm 1$ and C is a word on S.

Moreover, if $e = -1$, then C is equal in E to a word on $\{A_j : j \in J(v)\}$ (i.e., C is equal in E to an element of $A(v) \subset A$); if $e = +1$, then C is equal in E to a word in $\{B_j : j \in J(v)\}$ (i.e., C is equal in E to an element of $B(v) \subset B$).

REMARK A word of the form $p_v^e C p_v^{-e}$ is called a **pinch** if $e = \pm 1$, C is a word on S, and the "Moreover" paragraph above holds for C. The conclusion of Britton's lemma is that the word W in the hypothesis must contain a pinch as a subword.

Proof We first prove the special case of only one stable letter $p_v = p$. In this case, $J(v) = J$ and the subgroups $A(v)$ and $B(v)$ coincide with A and B, respectively. Thus, the presentation we deal with is

$$E^* = (S, p \mid D, p^{-1}A_j p = B_j, \quad j \in J).$$

We are done if W contains a subword of the form $p^{-1}p$ or pp^{-1}, so that we may assume

$$W \equiv W_0 p^{a_1} W_1 \cdots p^{a_n} W_n,$$

where $n \geq 1$, each a_i is a nonzero integer, and the W_i are words on S of which only W_0 and W_n may be empty.

We wish to use Lemma 12.6, but it does not apply because A_j and B_j may be distinct. Since the isomorphism condition holds, Lemma 12.7 gives $E \leq E^*$. If H is the group with presentation

$$H = (S, q \mid D, q^{-1}A_j q = B_j, \quad j \in J),$$

then clearly $H \cong E^*$. Consider now

$$H^* = (\{S, q\}, t \mid \{D, q^{-1}A_j q = B_j, \quad j \in J\}, \quad t^{-1}B_j t = B_j, \quad j \in J)$$

and the word

$$U \equiv W_0 (qt)^{a_1} W_1 \cdots (qt)^{a_n} W_n.$$

First of all, we claim that $U = 1$ in H^*. To see this, add the new generator p and the defining relation $p = qt$ to H^* to obtain a group K, where

$$K = (S, q, t, p \mid D, q^{-1}A_j q = B_j, \quad p = qt, \quad p^{-1}A_j p = B_j, \quad j \in J).$$

Now $K \cong H^*$ and K contains all the generators and relations of E^*.

Therefore, $W = 1$ in E^* implies $U = 1$ in H^*. Applying Lemma 12.6 to the groups H and H^* and the word U, we see that U contains a subword $t^eCt^{-e}(e = \pm 1)$, where C belongs to the subgroup B generated by the B_j. If $e = 1$, then $C \equiv W_k$ for some k ($1 \leq k \leq n - 1$), and we are done. If $e = -1$, then $C \equiv q^{-1}W_kq$ for some k; therefore, $q^{-1}W_kq \in B$ and so $W_k \in qBq^{-1} = A$.

We now consider the general case. Denote the elements of the index set V by 1, 2, $\cdots$. Let $D(v)$ consist of the relations $p_{v_j}^{-1}A_jp_{v_j} = B_j$ in E^* such that $j \in J(v)$. Since $W = 1$ in E^*, we have $W = 1$ in E_r for some r, where

$$E_r = (S, p_1, \cdots, p_r \mid D, D(1), \cdots, D(r)).$$

It is straightforward to show that $E \leq E_1 \leq E_2 \leq \cdots$. Choose s maximal such that W involves p_s; the chain of inequalities implies that $W = 1$ in E_s. Since E_s has a basis E_{s-1} and stable letter p_s, it follows from the first portion of our proof that W contains a subword $p_s^eCp_s^{-e}$, where, for example, $e = 1$ and C belongs to $B(s)$, the subgroup of E_{s-1} (and hence of E) generated by all B_j, $j \in J(s)$. (A similar conclusion holds if $e = -1$.) If the word C is a word on S, we are done. If, on the other hand, the word C involves some of $p_1, \cdots, p_{s-1}$, then there is a word C' on S (namely, a product of $B_j^{\pm 1}$, $j \in J(s)$) such that $C = C'$ in E_{s-1}. Thus, $CC'^{-1} = 1$ in E_{s-1}, and an induction on s implies that the word CC'^{-1}, hence C, hence W, contains a subword of the desired type. ∎

Definition Given E^* with basis E and stable letters $P = \{p_v : v \in V\}$, then a word W is **p_v-reduced** (for some fixed p_v) if W contains no pinch $p_v^eCp_v^{-e}$ as a subword. We say that W is ***P*-reduced** if W is p_v-reduced for all $p_v \in P$.

It is easy to see that an arbitrary word W is equal in E^* to a P-reduced word; the relations of E^* allow us to replace each pinch by a subword involving two fewer occurrences of a stable letter.

Corollary 12.9 *Let E^* be a Britton extension of E with stable letters $P = \{p_v : v \in V\}$. Assume $U \equiv R_0p_v^{e_1}R_1 \cdots p_v^{e_m}R_m$ and $W \equiv L_0p_v^{f_1}L_1 \cdots p_v^{f_n}L_n$ are P-reduced words, where each $e, f = \pm 1$, and none of the (possibly empty) words R or L involve p_v.*

If $U = W$ in E^*, then $m = n$ and $(e_1, \cdots, e_m) = (f_1, \cdots, f_n)$; moreover, the word $p_v^{e_m}R_mL_n^{-1}p_v^{-f_n}$ is a pinch.

Proof Since $UW^{-1} = 1$ in E^*, Britton's lemma asserts that it contains a pinch; since U and W (hence W^{-1}) are p_v-reduced, the pinch must occur at the interface: the subword $p_v^{e_m} R_m L_n^{-1} p_v^{-f_n}$ is a pinch. In particular, e_m and $-f_n$ must have opposite sign, so that $e_m = f_n$.

The remainder of the corollary is proved by induction on $\max\{m, n\}$. In the pinch $p_v^e R_m L_n^{-1} p_v^{-e}$ displayed above, we know from Britton's lemma that

$$R_m L_n^{-1} = \begin{cases} A_1^{\alpha_1} \cdots A_t^{\alpha_t} & \text{if } e = -1 \\ B_1^{\alpha_1} \cdots B_t^{\alpha_t} & \text{if } e = +1, \end{cases}$$

where $\alpha_i = \pm 1$, $A_i \in A(v)$, and $B_i \in B(v)$. In the first case,

$$\begin{aligned} p_v^{-1} R_m L_n^{-1} p_v &= p_v^{-1} A_1^{\alpha_1} \cdots A_t^{\alpha_t} p_v \\ &= (p_v^{-1} A_1^{\alpha_1} p_v)(p_v^{-1} \cdots p_v)(p_v^{-1} A_t^{\alpha_t} p_v) \\ &= B_1^{\alpha_1} \cdots B_t^{\alpha_t}. \end{aligned}$$

We have eliminated one p_v from U and one from W, so the remainder follows from the inductive hypothesis. The proof for the case $e = +1$ is similar. ∎

We shall now give a nice application of Britton's lemma. First, we need an easy remark.

Lemma 12.10 *Suppose G is a group having a presentation*

$$G_1 = (a_1, \cdots, a_m \mid r_n = 1, \quad n \in \mathbf{N})$$

as well as a finite presentation

$$G_2 = (b_1, \cdots, b_k \mid s_1 = 1, \cdots, s_t = 1).$$

Then all but a finite number of the relations $r_n = 1$ in the first presentation are superfluous.

Proof Let $\varphi: G_1 \to G_2$ be an isomorphism with inverse $\psi: G_2 \to G_1$. For each i, $\psi(s_i) = 1$ in G_1, and hence is a word on conjugates of various r_n. Since words have finite length and since there are only finitely many s_i, only a finite number of r_n are needed to prove $\psi(s_i) = 1$ in G_1 for all i. For notational convenience, let us denote these $r_1, r_2, \cdots, r_N$.

For each n, $\varphi(r_n) = 1$ in G_2 so that

$$\varphi(r_n) = w_n(s_1, \cdots, s_t),$$

a word on conjugates of $s_1, \cdots, s_t$. Therefore

$$r_n = \psi\varphi(r_n) = \psi(w_n(s_1, \cdots, s_t)) = w_n(\psi(s_1), \cdots, \psi(s_t)).$$

This equation says that r_n lies in the normal subgroup of G_1 generated by $r_1, r_2, \cdots, r_N$, which is what was to be proved. ∎

As we mentioned at the beginning of this chapter, the following theorem was first proved by B. H. Neumann, who observed that there are uncountably many nonisomorphic finitely generated groups but (obviously) only countably many finitely presented groups. The following explicit example is due to W. W. Boone.

Theorem 12.11 *There exists a finitely generated group that is not finitely presented.*

Proof Let $\langle a, b\rangle$ be a free group of rank 2. By Theorem 11.27, $\langle a, b\rangle$ contains a free subgroup of infinite rank, say, with free generating set $\{w_1, w_2, \cdots, w_n, \cdots\}$. Define a group G by the presentation

$$G = (a, b, p \mid p^{-1}w_n p = w_n, \quad n = 1, 2, \cdots).$$

This presentation exhibits G as a Britton extension of $\langle a, b\rangle$, as in Lemma 12.6.

Suppose G were finitely presented. By Lemma 12.10, we may delete all but a finite number of the relations. For notational convenience, we may assume

$$G = (a, b, p \mid p^{-1}w_n p = w_n, \quad n = 1, 2, \cdots, M - 1).$$

Now $p^{-1}w_M p = w_M$ in G, so that Britton's lemma provides a word Y in $\langle w_1, \cdots, w_{M-1}\rangle$ with

$$w_M = Y \quad \text{in} \quad \langle a, b\rangle.$$

This contradicts the fact that $\{w_1, w_2, \cdots, w_M\}$ freely generates its subgroup. ∎

EXERCISE

12.18. Use Britton's lemma to give a sharper form of Theorem 11.36: If A and B are subgroups of G and $\varphi: A \to B$ is an isomorphism,

then $G \leq (S, t \mid D, t^{-1}at = \varphi(a), a \in A)$, where $G = (S \mid D)$ is a presentation of G.

THE NOVIKOV-BOONE THEOREM

This section contains the proof of the following result.

Theorem 12.12 (Novikov-Boone) *There exists a finitely presented group G having an unsolvable word problem.*

We assure the reader that all the mathematical logic required in the proof has already been used; we need only Corollary 12.5, a paraphrase of Post's theorem, that exhibits a particular finitely presented semigroup Γ with an unsolvable word problem. Recall that Γ has a presentation

$$\Gamma = (q, q_1, \cdots, q_N, s_1, \cdots, s_M \mid F_i q_{i_1} G_i = H_i q_{i_2} K_i, \quad i \in I),$$

where F_i, G_i, H_i, K_i are positive words on $S = \{s_b : b = 1, \cdots, M\}$, and $q_{i_1}, q_{i_2} \in \{q, q_1, \cdots, q_N\}$.

The following notation is convenient. If $X \equiv s_1^{\varepsilon_1} \cdots s_m^{\varepsilon_m}$ is a (not necessarily positive) word on S, then $\overline{X} \equiv s_1^{-\varepsilon_1} \cdots s_m^{-\varepsilon_m}$. If X and Y are words on S, then

$$(Xq_jY)^* \equiv \overline{X} q_j Y,$$

where $q_j \in \{q, q_1, \cdots, q_N\}$.

Definition A word Σ is **special** if $\Sigma \equiv \overline{X} q_j Y$, where X and Y are positive words on S, and $q_j \in \{q, q_1, \cdots, q_N\}$.

We now present a group G that will be seen to have an unsolvable word problem. It has generators

$$q, q_1, \cdots, q_N, s_1, \cdots, s_M, \quad r_i \, (i \in I), x, t, k,$$

and relations, for all $i \in I$ and $b = 1, \cdots, M$,

$$\begin{aligned}
xs_b &= s_b x^2 && \Big] && \Big] && \Big]\Delta_1 \\
r_i s_b &= s_b x r_i x && && \Big]\Delta_2 \\
r_i^{-1} \overline{F}_i q_{i_1} G_i r_i &= \overline{H}_i q_{i_2} K_i && && \\
tr_i &= r_i t, \quad tx = xt && \Big]\Delta_3 && \\
kr_i &= r_i k, \quad kx = xk && && \\
k(q^{-1}tq) &= (q^{-1}tq)k. && &&
\end{aligned}$$

The subsets $\Delta_1 \subset \Delta_2 \subset \Delta_3$ of the relations are labelled for future reference.

The reduction to Post's theorem is accomplished by the following lemma.

Lemma 12.13 (Boone) *If Σ is a special word, then*

$$k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k \quad \text{in} \quad G \tag{1}$$

if and only if $\Sigma^* = q$ in Γ.

There is an immediate proof of the Novikov-Boone theorem from Boone's lemma: If there were a decision process to determine whether two words are equal in G, then there would be a decision process to determine, for an arbitrary special word Σ, whether $\Sigma^* = q$ in Γ. But Corollary 12.5 asserts that no such decision process for Γ exists.

Proof of Sufficiency in Boone's Lemma Assume that $\Sigma^* \equiv Xq_jY = q$ in Γ, where X and Y are positive words on $S = \{s_b : b = 1, \cdots, M\}$ and $1 \leq j \leq N$.

First, we claim that if V is a positive word on S, then

$$r_iV = VR \quad \text{in} \quad G,$$

where R is a positive word on r_i and x. We prove this by induction on m, where $V = s_{b_1} \cdots s_{b_m}$. The claim is certainly true when $m = 0$. If $m > 0$, then $V = V_1 s_{b_m}$, where $V_1 = s_{b_1} \cdots s_{b_{m-1}}$. By induction, $r_iV = V_1R_1s_{b_m}$, where R_1 is a positive word in r_i and x. (A similar argument shows that $r_i^{-1}V = VR'$ in G.) Using the relations $xs_b = s_bx^2$ and $r_is_b = s_bxr_ix$, we see that $R_1s_{b_m} = s_{b_m}R$ for some positive word R in r_i and x.

It follows easily that if U is a positive word on S, then

$$\overline{U}r_i^{-1} = L\overline{U} \quad \text{in} \quad G,$$

where L is a word on r_i and x: If $U \equiv s_{b_1} \cdots s_{b_m}$, set $V \equiv s_{b_m} \cdots s_{b_1}$ (so that $\overline{U} = V^{-1}$); then $r_iV = VR$ (as above) and

$$\overline{U}r_i^{-1} = V^{-1}r_i^{-1} = (r_iV)^{-1} = (VR)^{-1} = R^{-1}V^{-1} = R^{-1}\overline{U}.$$

A similar argument shows that $\overline{U}r_i = L'\overline{U}$, where L' is a word on r_i, x.

Assume now that Σ is a special word with $\Sigma^* = q$ in Γ. There is a sequence of elementary operations in Γ

$$\Sigma^* \equiv W_1 \to W_2 \to \cdots \to W_n \equiv q$$

such that, for each v, one of the words W_v, W_{v+1} has the form $UF_iq_{i_1}G_iV$, the other $UH_iq_{i_2}K_iV$, where U and V are positive words on S. Now, in G we have equations

$$\begin{aligned}\overline{U}(\overline{H}_iq_{i_2}K_i)V &= \overline{U}(r_i^{-1}\overline{F}_iq_{i_1}G_ir_i)V \\ &= L'\overline{U}(\overline{F}_iq_{i_1}G_i)VR',\end{aligned}$$

where L' and R' are words in r_i and x. (In a similar manner, one sees that $\overline{U}(\overline{F}_iq_{i_1}G_i)V = \overline{U}(r_i\overline{H}_iq_{i_2}K_ir_i^{-1})V = L''\overline{U}(\overline{H}_iq_{i_2}K_i)VR''$ in G.) Since $W_v = W_{v+1}$ in Γ implies $W_v^* = W_{v+1}^*$ in G, it follows that, for each v,

$$W_v^* = L_vW_{v+1}^*R_v \quad \text{in} \quad G,$$

where L_v and R_v are words in x and r_i. Setting $L = L_1L_2\cdots L_{n-1}$ and $R = R_{n-1}\cdots R_2R_1$, we see that

$$W_1^* = LW_n^*R \quad \text{in} \quad G.$$

But $W_1^* = (\Sigma^*)^* = \Sigma$ and $W_n^* = q^* = q$. Therefore, if Σ is a special word with $\Sigma^* = q$ in Γ, then

$$\Sigma = LqR \quad \text{in} \quad G,$$

where L and R are words in x and various r_i.

Let us now prove equation (1). We have just seen that $\Sigma = LqR$ in G, where L and R both commute with t and k (since x and all r_i commute with t and k). Thus,

$$\begin{aligned}(\Sigma^{-1}t\Sigma)k &= R^{-1}q^{-1}(L^{-1}tL)q(Rk) \\ &= R^{-1}[(q^{-1}tq)k]R \\ &= (R^{-1}k)q^{-1}tqR \\ &= kR^{-1}q^{-1}L^{-1}tLqR \\ &= k(\Sigma^{-1}t\Sigma). \quad \blacksquare\end{aligned}$$

Proof of Necessity of Boone's Lemma The proof of necessity of Boone's lemma will be in several steps. First of all, we show that the group G is the end result of a chain of Britton extensions. Define groups G_0, G_1, G_2, and G_3 as follows:

$G_0 = \langle x \rangle$, the free group on x;

$G_1 = (G_0\,;\, s_b, b = 1, \cdots, M \mid \text{relations } \Delta_1 \text{ indicated above})$

(this notation means we are adjoining the displayed generators and relations to the given presentation of G_0);

$$G_2 = (G_1 ; r_i, \quad i \in I, \quad q, q_1, \cdots, q_N \mid \text{relations } \Delta_2);$$

$$G_3 = (G_2 ; t \mid \text{relations } \Delta_3).$$

Lemma 12.14 *G is a Britton extension of G_3 with stable letter k; G_3 is a Britton extension of G_2 with stable letter t.*

Proof First note that G has basis G_3 and stable letter k:

$$G = (\_\_, k \mid \_\_, k^{-1}r_ik = r_i, k^{-1}xk = x, k^{-1}(q^{-1}tq)k = q^{-1}tq);$$

also, G_3 has basis G_2 and stable letter t:

$$G_3 = (\_\_, t \mid \_\_, t^{-1}r_it = r_i, t^{-1}xt = x).$$

Each of these presentations is of the type occurring in Lemma 12.6, so that each is a Britton extension, i.e., the isomorphism condition holds. ∎

Lemma 12.15 *G_2 is a Britton extension of $G_1 * \langle q, q_1, \cdots, q_N\rangle$ with stable letters $\{r_i, i \in I\}$; G_1 is a Britton extension of G_0 with stable letters $\{s_b, \quad b = 1, \cdots, M\}$.*

Proof Note that G_2 has basis $G_1 * \langle q, q_1, \cdots, q_N\rangle$ and stable letters $\{r_i : i \in I\}$:

$$G_2 = (\_\_, \{r_i : i \in I\} \mid \_\_, r_i^{-1}(\overline{F}_iq_{i_1}G_i)r_i = \overline{H}_iq_{i_2}K_i, r_i^{-1}(s_bx)r_i = s_bx^{-1}).$$

To verify the isomorphism condition, first observe that for each index i, the subgroup $A(i) = \langle \overline{F}_iq_{i_1}G_i, s_bx, \ b = 1, \cdots, M\rangle$ and the subgroup $B(i) = \langle \overline{H}_iq_{i_2}K_i, s_bx^{-1}, \ b = 1, \cdots, M\rangle$. It is easy to check that $A(i)$ and $B(i)$ are each free on the displayed generating sets (for example, map $A(i)$ onto the free group with free basis $\{q_{i_1}, s_1, \cdots, s_M\}$ by setting x equal to 1 and using Exercise 12.16). There is thus an isomorphism $\varphi_i : A(i) \to B(i)$ with $\varphi_i(\overline{F}_iq_{i_1}G_i) = \overline{H}_iq_{i_2}K_i$ and $\varphi_i(s_bx) = s_bx^{-1}$, $b = 1, \cdots, M$.

The presentation

$$G_1 = (x, s_1, \cdots, s_M \mid s_b^{-1}xs_b = x^2, \quad b = 1, \cdots, M)$$

shows that G_1 has basis $\langle x\rangle$ and stable letters $s_1, \cdots, s_M$. Since x has infinite order, $\langle x\rangle \cong \langle x^2\rangle$, so that the isomorphism condition holds. ∎

We remark that the fact that $\{s_1x, \cdots, s_Mx\}$ freely generates its subgroup (just proved above) will be used several times in the sequel.

EXERCISE

**12.19. There is an automorphism ψ of G_1 sending x to x^{-1} and fixing each s_b.

Lemma 12.16 *Let Σ be a fixed special word satisfying the hypothesis of Boone's lemma:*

$$k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k \quad in \quad G.$$

Then there are words L_1 and L_2 on x and various r_i such that

$$L_1\Sigma L_2 = q \quad in \quad G_2.$$

Proof By hypothesis,

$$k^{-1}\Sigma^{-1}t\Sigma k\Sigma^{-1}t^{-1}\Sigma = 1 \quad \text{in} \quad G.$$

Since G is a Britton extension of G_3 with stable letter k, we have by Britton's lemma (actually, by the simpler Lemma 12.6)

$$\Sigma^{-1}t\Sigma = C' \quad \text{in} \quad G_3,$$

where C' is a word in various r_i, x, and $q^{-1}tq$. Therefore

$$W \equiv \Sigma^{-1}t\Sigma R_0(q^{-1}t^{e_1}q)R_1(q^{-1}t^{e_2}q)R_2 \cdots q^{-1}t^{e_n}qR_n = 1 \quad \text{in} \quad G_3,$$

where the R_j are (possibly empty) words on x and various r_i, and the exponents $e_j = \pm 1$. We assume this equation is chosen so that n is minimal.

By Britton's lemma (really, by Lemma 12.6 again), W contains a pinch t^eCt^{-e}, where

$$C = R \quad \text{in} \quad G_2$$

for some word R in the r_i and x. Note that $C = R$ in G_3 since $G_2 \leq G_3$.

If the initial letter t^e is, in fact, the first t occurring in W, then $e = 1$ and $t^eCt^{-e} \equiv t\Sigma R_0q^{-1}t^{e_1}$, for C does not involve t. Hence

$$\Sigma R_0q^{-1} \equiv C = R \quad \text{in} \quad G_2$$

or, equivalently,

$$R^{-1}\Sigma R_0 = q \quad \text{in} \quad G_2,$$

which is of the desired form.

If the initial t^e in t^eCt^{-e} is t^{e_j} for some $j \geq 1$, then we have

$$t^eCt^{-e} \equiv t^{e_j}qR_jq^{-1}t^{e_{j+1}},$$

for C does not involve t. Thus, $e_j = e$ and $e_{j+1} = -e$. Hence, since t commutes with x and the r_i, we have in G_3:

$$\begin{aligned} q^{-1}t^{e_j}qR_jq^{-1}t^{e_{j+1}}q &\equiv q^{-1}t^eCt^{-e}q \\ &= q^{-1}t^eRt^{-e}q \\ &= q^{-1}Rq \\ &= q^{-1}Cq \\ &\equiv q^{-1}(qR_jq^{-1})q = R_j. \end{aligned}$$

Therefore, W has a factorization in G_3 of shorter length, thus contradicting the choice of minimal n. ▮

If we recall that $\Sigma \equiv \bar{X}q_jY$, then we have just shown that

$$L_1\bar{X}q_jYL_2 = q \quad \text{in} \quad G_2$$

for some words L_1 and L_2 on x and various r_i. We rewrite this last equation

$$L_1\bar{X}q_j = qL_2^{-1}Y^{-1} \quad \text{in} \quad G_2.$$

Definition A word w on letters A is **freely reduced** if it contains no subwords of the form aa^{-1} or $a^{-1}a$, where $a \in A$.

Clearly we may assume the words L_1 and L_2 above are freely reduced (if there were such a subword, cancel it, and the equation still holds).

Lemma 12.17 *If L_1 and L_2 are freely reduced, then each of the words $L_1\bar{X}q_j$ and $qL_2^{-1}Y^{-1}$ are r_i-reduced for all i.*

Proof Let us show that $L_1\bar{X}q_j$ contains no pinch of the form $r_i^eCr_i^{-e}$ (the proof for the other word is similar). In view of the spelling of $L_1\bar{X}q_j$, any such pinch is a subword of L_1, so that $C \equiv x^m$ for some $m \neq 0$ (since L_1 is freely reduced). We must show that x^m is not equal in $G_1 * \langle q, q_1, \cdots, q_N\rangle$ to a word on

$$\{\bar{F}_iq_{i_1}G_i, s_1x, \cdots, s_Mx\} \quad \text{or on} \quad \{\bar{H}_iq_{i_2}K_i, s_1x^{-1}, \cdots, s_Mx^{-1}\}.$$

Suppose $x^m = V$ in $G_1 * \langle q, q_1, \cdots, q_N\rangle$, where

$$V \equiv w_0(\bar{F}_iq_{i_1}G_i)^{e_1} \cdots (\bar{F}_iq_{i_1}G_i)^{e_n}w_n$$

and each $e_j = \pm 1$, w_j is a word on $\{s_1 x, \cdots, s_M x\}$, and V is reduced as an element in a free product. Since $x^m \in G_1$, one of the free factors, the word V cannot involve q_{i_1}, hence has no occurrence of $\bar{F}_i q_{i_1} G_i$. Therefore

$$x^m = w_0 \quad \text{in} \quad G_1.$$

Write $w_0 = (s_{b_1} x)^{f_1} \cdots (s_{b_K} x)^{f_K}$, where each $f_\nu = \pm 1$. Since $x^{-m} w_0 = 1$ in G_1, which is a Britton extension of $\langle x \rangle$, the word $x^{-m} w_0$ must contain a pinch $s_b^f x^\varepsilon s_b^{-f}$, where $\varepsilon = \pm 1$. Britton's lemma says further that if $f = +1$, then $x^\varepsilon \in \langle x^2 \rangle$, which is impossible. Thus the pinch is $s_b^{-1} x^\varepsilon s_b$. Inspection of the spelling of w_0, however, shows that it can contain no such subword. Therefore $w_0 = 1$ and $x^m = 1$. Since x has infinite order, $m = 0$, contradicting our earlier hypothesis that $m \neq 0$.

A similar argument shows that x^m is not equal in

$$G_1 * \langle q, q_1, \cdots, q_N \rangle$$

to a word on $\{\bar{H}_i q_{i_2} K_i, s_1 x^{-1}, \cdots, s_M x^{-1}\}$. ∎

We now know that $L_1 \bar{X} q_j = q L_2^{-1} Y^{-1}$ in G_2 and that each word is r_i-reduced for each i. As no r_i-letters occur outside of L_1 or L_2, Corollary 12.9 says that the number T of various r_i-letters in L_1 is the same as in L_2.

The following lemma completes the proof of Boone's lemma and, with it, the Novikov-Boone theorem. In view of a further application in the next section, we prove slightly more than we need now.

Lemma 12.18 *Let L_1 and L_2 be words on various r_i and x that are r_i-reduced, each r_i. If X and Y are freely reduced words on $S = \{s_1, \cdots, s_M\}$ and if*

$$L_1 \bar{X} q_j Y L_2 = q \quad \text{in} \quad G_2,$$

then both X and Y are positive words and

$$(\bar{X} q_j Y)^* = q \quad \text{in} \quad \Gamma.$$

REMARK The hypothesis that X and Y be freely reduced words on S is automatically true in our case, for we have X and Y positive at the outset.

Proof The proof is by induction on the number T of r_i-letters occurring in L_1 (which is the same as for L_2, by our remarks above). If $T = 0$, the equation $L_1 \bar{X} q_j Y L_2 = q$ in G_2 is just

$$x^m \bar{X} q_j Y x^n = q \quad \text{in} \quad G_2.$$

Since no r_i appears and since $G_1 * \langle q, q_1, \cdots, q_N\rangle \leq G_2$, we have

$$x^m\bar{X}q_jYx^n = q \quad \text{in} \quad G_1 * \langle q, q_1, \cdots, q_N\rangle.$$

Such an equation can hold in a free product only if $q_j = q$ and $x^m\bar{X} = 1 = Yx^n$ in G_1. A glance at the presentation of G_1 shows $m = n = 0$ and X and Y are empty. Thus, we surely have X and Y positive and $Xq_jY = q$ in Γ.

Assume now that $T > 0$. By Corollary 12.9, we may write

$$L_1\bar{X}q_jYL_2 \equiv L_3[r_i^e x^m\bar{X}q_jYx^n r_i^{-e}]L_4 = q \quad \text{in} \quad G_2,$$

where the word in brackets is a pinch. By Britton's lemma, the element $x^m\bar{X}q_jYx^n$ lies in the subgroup

$$A(i) = \langle \bar{F}_i q_{i_1} G_i, s_1x, \cdots, s_Mx\rangle \qquad \text{if} \qquad e = -1,$$

and in the subgroup

$$B(i) = \langle \bar{H}_i q_{i_2} K_i, s_1x^{-1}, \cdots, s_Mx^{-1}\rangle \qquad \text{if} \qquad e = +1.$$

In the first case, $j = i_1$, i.e.,

$$q_j = q_{i_1}; \tag{2}$$

in the second case, $j = i_2$, i.e., $q_j = q_{i_2}$ (because, for example, $x^m\bar{X}q_jYx^n \in A(i)$ is valid in the free product $G_1 * \langle q, q_1, \cdots, q_N\rangle$). Let us focus on the case $e = -1$, the case $e = +1$ being similar. We have

$$\begin{aligned} W &\equiv x^m\bar{X}q_jYx^n u_0(\bar{F}_i q_{i_1} G_i)^{\alpha_1} u_1 \cdots (\bar{F}_i q_{i_1} G_i)^{\alpha_t} u_t \\ &= 1 \quad \text{in} \quad G_1 * \langle q, q_1, \cdots, q_N\rangle, \end{aligned}$$

where $\alpha_j = \pm 1$ and the u_j are (possibly empty) words on

$$\{s_1x, \cdots, s_Mx\}.$$

Assume that t is minimal and each u_j is reduced as a word on the free generators $\{s_1x, \cdots, s_Mx\}$. Since $G_1 * \langle q, q_1, \cdots, q_N\rangle$ is a Britton extension of G_1 and W involves a stable letter (one of the q's), Britton's lemma asserts that W contains a pinch. If a pinch involves the first occurrence of q_j, then $\alpha_1 = -1$ and $Yx^nu_0G_i^{-1} = 1$ in G_1. We claim that a pinch cannot occur elsewhere in W. Suppose a pinch occurs as a subword of $(\bar{F}_i q_{i_1} G_i)^{\alpha_j} u_j (\bar{F}_i q_{i_1} G_i)^{\alpha_{j+1}}$. If $\alpha_j = +1$, then $\alpha_{j+1} = -1$ and $G_i u_j G_i^{-1} = 1$ in G_1; if $\alpha_j = -1$, then $\alpha_{j+1} = +1$ and $\bar{F}_i^{-1} u_j \bar{F}_i = 1$ in G_1. In either case, we have $u_j = 1$ in G_1.

But u_j is a reduced word on the free generators $\{s_1x, \cdots, s_Mx\}$; therefore $u_j \equiv 1$, which violates the minimality of t. We conclude that $t = 1$, $\alpha_1 = -1$ and

$$W \equiv x^m\overline{X}q_jYx^nu_0G_i^{-1}q_{i_1}^{-1}\overline{F}_i^{-1}u_1 = 1 \quad \text{in} \quad G_1 * \langle q, q_1, \cdots, q_N\rangle.$$

We have already seen that

$$Yx^nu_0G_i^{-1} = 1 \quad \text{in} \quad G_1.$$

It follows from W being a word in a free product that

$$\overline{F}_i^{-1}u_1x^m\overline{X} = 1 \quad \text{in} \quad G_1.$$

We recast these equations into more convenient form (by conjugating):

$$x^nu_0G_i^{-1}Y = 1 \quad \text{in} \quad G_1$$

and

$$\overline{X}\overline{F}_i^{-1}u_1x^m = 1 \quad \text{in} \quad G_1.$$

Since u_0 and u_1 are words on free generators $\{s_1x, \cdots, s_Mx\}$ (proof of Lemma 12.15) we may assume they are reduced; it follows that we may assume u_0 and u_1 contain no subwords of the form $s_bs_b^{-1}$ or $s_b^{-1}s_b$.

We now claim that after cancelling all subwords (if any) of the form $s_bs_b^{-1}$ or $s_b^{-1}s_b$, the first surviving letter in $G_i^{-1}Y$ is positive (in particular, the whole of G_i^{-1} disappears). Otherwise, the word $G_i^{-1}Y$ begins with s_b^{-1} for some b. Since

$$x^nu_0G_i^{-1}Y = 1 \quad \text{in} \quad G_1$$

is a word involving s_b, Britton's lemma provides a pinch $s_b^eCs_b^{-e}$. As u_0 is written as a reduced word in the free group on $\{s_1x, \cdots, s_Mx\}$, the pinch is not a subword of x^nu_0. It follows that the pinch $s_b^eCs_b^{-e}$ must have, as its last letter, the first surviving letter in $G_i^{-1}Y$. If this letter is not positive, as we are assuming, then $e = 1$ and $s_b^eCs_b^{-e} \equiv s_bxs_b^{-1}$ (for C is a word on x). But we may further conclude from Britton's lemma that $x \in \langle x^2\rangle$, and this is a contradiction.

In a similar manner, one sees that after cancellation of all subwords of the form $s_bs_b^{-1}$ or $s_b^{-1}s_b$ (if any) in $\overline{X}\overline{F}_i^{-1}$, the last surviving letter is "negative" (in particular, the whole of $\overline{F}_i^{-1}$ disappears). The proof is just as above, inverting the original equation $\overline{X}\overline{F}_i^{-1}u_1x^m = 1$ in G_1.

We have proved that

$$u_0^{-1} = (G_i^{-1}Y)x^n \quad \text{in} \quad G_1,$$

where $G_i^{-1}Y \equiv Y_1$ is a positive word on S and, by definition, u_0^{-1} is a word on $s_1x, \cdots, s_Mx$. Define $v_0^{-1} = r_i^{-1}u_0^{-1}r_i$. Since $r_i^{-1}s_bxr_i = s_bx^{-1}$, all b, the element v_0^{-1} is a word in $s_1x^{-1}, \cdots, s_Mx^{-1}$. But we may also regard u_0^{-1} and v_0^{-1} as elements of $G_1 = \langle x, s_1, \cdots, s_M\rangle$. By Exercise 12.19, there is an automorphism ψ of G_1 taking x to x^{-1} and fixing all s_b; thus, $\psi(u_0^{-1}) = v_0^{-1}$. It follows that

$$v_0^{-1} = Y_1x^{-n} \quad \text{in} \quad G_1,$$

where Y_1 is a positive word on S. A similar argument gives

$$v_1^{-1} = x^{-m}\overline{X}_1 \quad \text{in} \quad G_1,$$

where $\overline{X}_1 \equiv \overline{X}\overline{F}_i^{-1}$ is a "negative" word on S, and $r_i^{-1}u_1^{-1} = v_1^{-1}r_i^{-1}$. Let us return to the induction. We have in G_2,

$$\begin{aligned} q = L_1\overline{X}q_jYL_2 &\equiv L_3r_i^{-1}(x^m\overline{X})q_j(Yx^n)r_iL_4 \\ &= L_3r_i^{-1}(u_1^{-1}\overline{F}_i)q_j(G_iu_0^{-1})r_iL_4 \\ &= L_3v_1^{-1}r_i^{-1}(\overline{F}_iq_jG_i)r_iv_0^{-1}L_4 \\ &= L_3v_1^{-1}(r_i^{-1}\overline{F}_iq_{i_1}G_ir_i)v_0^{-1}L_4 \\ &= (L_3v_1^{-1})\overline{H}_iq_{i_2}K_i(v_0^{-1}L_4) \\ &= (L_3x^{-m}\overline{X}_1)\overline{H}_iq_{i_2}K_i(Y_1x^{-n}L_4). \end{aligned}$$

We have verified that

$$L_3x^{-m}(\overline{X}_1\overline{H}_iq_{i_2}K_iY_1)x^{-n}L_4 = q \quad \text{in} \quad G_2.$$

Now L_3x^{-m} and $x^{-n}L_4$ are words in x and various r_i having at most $T-1$ occurrences of various r_i. In order to apply the inductive hypothesis, we must check that $\overline{X}_1\overline{H}_i$ and K_iY_1 are freely reduced, i.e., contain no subwords of form $s_bs_b^{-1}$ or $s_b^{-1}s_b$. Now K_i is positive, hence contains no "forbidden" subwords; further, $Y_1 \equiv G_i^{-1}Y$ is just a subword of Y (since G_i^{-1} disappears), and hence has no forbidden subwords, by hypothesis. The only place where a forbidden subword can occur in K_iY_1 is at the interface; this is impossible, for we have seen that Y_1 begins with a positive letter. A similar argument shows that $\overline{X}_1\overline{H}_i$ is freely reduced.

By induction, both X_1H_i and K_iY_1 are positive; hence their subwords X_1 and Y_1 are also positive. Therefore $X \equiv X_1F_1$ and $Y \equiv K_iY_1$ are also positive. Induction also gives

$$(\overline{X}_1\overline{H}_iq_{i_2}K_iY_1)^* = q \quad \text{in} \quad \Gamma.$$

Therefore, since $\overline{\overline{X}_1\overline{H}_i} = X_1H_i$, we have

$$X_1H_iq_{i_2}K_iY_1 = q \quad \text{in} \quad \Gamma \tag{3}$$

(it is here we see why the bar operation is used; had we tried inversion instead, we would now only have $H_iX_1q_{i_2}K_iY_1 = q$ in Γ). Thus, using (2),

$$\begin{aligned} Xq_jY &\equiv X_1F_iq_jG_iY_1 \\ &\equiv X_1F_iq_{i_1}G_iY_1 = X_1H_iq_{i_2}K_iY_1 \quad \text{in} \quad \Gamma. \end{aligned}$$

Combining this with equation (3) gives

$$Xq_jY = q \quad \text{in} \quad \Gamma,$$

as desired. ∎

There are other proofs of the unsolvability of the word problem for groups, one due to P. Novikov, another due to G. Higman. Higman's proof is a corollary of his imbedding theorem, which we prove in the next section. The proof we shall give uses our development so far, whereas Higman's original proof of his imbedding theorem is independent of the material in this chapter.

There are other group-theoretical properties which yield unsolvable problems; the reader may find details in the book of C. F. Miller, III. On the positive side, we must mention a theorem (1932) of W. Magnus proving that every finitely generated group having a single defining relation does have a solvable word problem.

AN IMBEDDING THEOREM OF G. HIGMAN

When can a finitely generated group be imbedded in a finitely presented group? The answer to this purely group-theoretical question, discovered by G. Higman in 1961, reveals a harmonic interplay of group theory with mathematical logic. The proof we present here is due to S. Aanderaa and was found in 1970.

We need an elementary technical observation before we can state Higman's theorem.

Lemma 12.19 *Every finitely generated group G has a presentation*

$$G = (a_1, \cdots, a_m | R_i = 1, \quad i \in I)$$

in which each R_i is a positive word on $a_1, \cdots, a_m$.

Proof If $(b_1, \cdots, b_n \mid d_j = 1,\ j \in J)$ is a presentation of G, then a new presentation is

$$G = (b_1, \cdots, b_n, c_1, \cdots, c_n \mid D_j = 1,\quad j \in J,\quad b_k c_k = 1),$$

where D_j is obtained from d_j by replacing each occurrence of b_k^{-1} by c_k. ∎

Definition A group R is **recursively presented** if it has a presentation

$$R = (u_1, \cdots, u_m \mid w = 1,\quad w \in E),$$

where each w is a positive word on $u_1, \cdots, u_m$ and E is an r.e. set.

We insist that the defining relations be positive words, for, according to our exposition, only sets of positive words are allowed to be r.e. sets. Lemma 12.19 shows that this is really no restriction.

EXERCISES

12.20. If a finitely generated group G has a presentation whose relations form an r.e. set of positive words, then G has a presentation whose relations form a recursive set of positive words. (HINT: Assume the given presentation is

$$G = (u_1, \cdots, u_m \mid w_i = 1,\quad i = 1, 2, \cdots);$$

define a new presentation

$$G = (u_1, \cdots, u_m, y \mid y = 1,\quad y^i w_i = 1,\quad i = 1, 2, \cdots).)$$

***12.21.** Every finitely generated subgroup of a finitely presented group is recursively presented. (HINT: Consider all words that can be obtained from 1 by a finite number of elementary operations.)

Theorem 12.20 (G. Higman) *Every recursively presented group R can be imbedded in a finitely presented group.*

(A characterization of finitely generated subgroups of finitely presented groups is thus provided by Higman's theorem and Exercise 12.21.)

Proof Assume R has a presentation

$$R = (u_1, \cdots, u_m \mid w = 1, \quad w \in E),$$

where E is an r.e. set of positive words on $u_1, \cdots, u_m$. There is thus a Turing machine T enumerating E whose alphabet $\{s_0, \cdots, s_M\}$ contains $\{u_1, \cdots, u_m\}$; moreover, Exercise 12.14 allows us to assume T has stopping state q_0. By Lemma 12.3, there is a finitely presented semigroup $\gamma = \gamma(T)$ for which

$$w \in E \quad \text{if and only if} \quad hq_1wh = q \quad \text{in} \quad \gamma.$$

From the semigroup γ, we constructed a semigroup Γ (essentially by relabelling the letter h so that it became s_M). From Γ, we constructed a group G with generators

$$q, q_1, \cdots, q_N, s_1, \cdots, s_M, r_i, x, t, k$$

and relations Δ_2 of the last section and

$$tr_i = r_it, \qquad tx = xt$$
$$kr_i = r_ik, \qquad kx = xk$$
$$k(q^{-1}tq) = (q^{-1}tq)k.$$

It is now more convenient to return to the original notation, i.e., from Γ back to γ. Let us first see what we need.

Boone's lemma (Lemma 12.13) says that if Σ is a special word, then

$$\Sigma^* = q \quad \text{in } \Gamma \text{ if and only if} \quad k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k \quad \text{in} \quad G.$$

In particular, $\Sigma \equiv h^{-1}q_1wh$ is a special word. Therefore the following are equivalent for a positive word w on $s_0, \cdots, s_M$:

$$w \in E;$$
$$\Sigma^* \equiv hq_1wh = q \quad \text{in} \quad \Gamma;$$
$$k(\Sigma^{-1}t\Sigma) = (\Sigma^{-1}t\Sigma)k \quad \text{in} \quad G;$$
$$(4) \quad k(h^{-1}w^{-1}q_1^{-1}hth^{-1}q_1wh) = (h^{-1}w^{-1}q_1^{-1}hth^{-1}q_1wh)k \quad \text{in} \quad G.$$

In order to simplify equation (4), we introduce a new presentation G' of G using the defining equations

$$k_0 = hkh^{-1} \qquad \text{and} \qquad t_0 = q_1^{-1}hth^{-1}q_1.$$

Recall that G_2 has generators (in the notation of γ)

$$q, h, q_0, \cdots, q_N, s_0, \cdots, s_M, r_i, x$$

and relations Δ_2:

$$xs_b = s_b x^2, \qquad xh = hx^2$$

$$r_i s_b = s_b x r_i x, \qquad r_i h = h x r_i x,$$

$$r_i^{-1} \bar{F}_i q_{i_1} G_i r_i = \bar{H}_i q_{i_2} K_i.$$

Define

$$G_3' = (G_2; t_0 \mid t_0^{-1}(q_1^{-1} h r_i h^{-1} q_1) t_0 = q_1^{-1} h r_i h^{-1} q_1, \quad t_0^{-1}(q_1^{-1} h x h^{-1} q_1) t_0 = q_1^{-1} h x h^{-1} q_1).$$

Note that G_3' is another presentation of the group G_3'; this is quickly seen by replacing t_0 by its definition. Similarly, we define another presentation of G:

$$G' = (G_3'; k_0 \mid k_0^{-1}(h r_i h^{-1}) k_0 = h r_i h^{-1}, \; k_0^{-1}(h x h^{-1}) k_0 = h x h^{-1}, \; k_0^{-1}(h q^{-1} h^{-1} q_1 t_0 q_1^{-1} h q h^{-1}) k_0 = h q^{-1} h^{-1} q_1 t_0 q_1^{-1} h q h^{-1}). \blacksquare$$

Lemma 12.21 *G_3' is a Britton extension of G_2 with stable letter t_0; G' is a Britton extension of G_3' with stable letter k_0.*

Proof Exactly as the proof of Lemma 12.14. ∎

Lemma 12.22 *Let w be a positive word on $s_0, \cdots, s_M$. Then*

$$w \in E \quad \text{if and only if} \quad k_0(w^{-1} t_0 w) = (w^{-1} t_0 w) k_0 \quad \text{in } G'.$$

Proof Equation (4) above has this simpler form when we change t and k to t_0 and k_0. ∎

Define a new group that is the free product:

$$G_4 = G' * R.$$

We consider the presentation of G_4 given by the union of the generators and relations of G' and of R. Since this union must be disjoint, let us consider letters $\{a_1, \cdots, a_m\} \subset \{s_0, \cdots, s_M\}$ with $a_i \leftrightarrow u_i$. We thus consider the set E as positive words on $\{a_1, \cdots, a_m\}$.

Define groups G_5, G_6, and G_7 as follows:

$$G_5 = (G_4; b_1, \cdots, b_m \mid b_i^{-1} u_j b_i = u_j, \; b_i^{-1} a_j b_i = a_j, \quad b_i^{-1} k_0 b_i = k_0 u_i^{-1}, \quad \text{all } i, j = 1, \cdots, m)$$

$$G_6 = (G_5; d \mid d^{-1} k_0 d = k_0, \quad d^{-1} a_i b_i d = a_i, \quad i = 1, \cdots, m)$$

$$G_7 = (G_6; \sigma \mid \sigma^{-1} t_0 \sigma = t_0 d, \quad \sigma^{-1} k_0 \sigma = k_0, \quad \sigma^{-1} a_i \sigma = a_i, \quad i = 1, \cdots, m).$$

We prove Higman's theorem by showing first that each of these groups is a Britton extension of its predecessor. Lemma 12.7 then will give

$$R \subset G_4 \leq G_5 \leq G_6 \leq G_7.$$

The proof is completed by showing G_7 has a finite presentation.

Lemma 12.23 *The subgroups* $\langle a_1, \cdots, a_m, k_0 \rangle_4$ *and* $\langle a_1, \cdots, a_m, t_0 \rangle_4$ *of* G_4 *are free on the displayed generating sets.*

REMARK The subscript 4 indicates that the subgroups are computed in G_4, i.e., only the relations of G_4 are used.

Proof Recall our analysis of the group G'; it is the end result of a chain of Britton extensions

$$G_1 * \langle q, q_0, \cdots, q_N \rangle \leq G_2 \leq G_3' \leq G',$$

where G_1 has basis $\langle x \rangle$ and stable letters $h, s_0, \cdots, s_M$. Exercise 12.17 gives $\langle h, s_0, \cdots, s_M \rangle_1$ free on these generators, so that Lemma 12.7 gives $\langle h, s_0, \cdots, s_M \rangle_4$ free on these generators. Since $\{a_1, \cdots, a_m\} \subset \{s_0, \cdots, s_M\}$, we have $\langle a_1, \cdots, a_m \rangle_4$ free on $a_1, \cdots, a_m$.

Let us prove $\langle a_1, \cdots, a_m, k_0 \rangle_4$ is free on $a_1, \cdots, a_m, k_0$; a similar argument works for $\langle a_1, \cdots, a_m, t_0 \rangle_4$. Suppose

$$W \equiv C_0 k_0^{e_1} C_1 k_0^{e_2} \cdots C_{n-1} k_0^{e_n} = 1 \quad \text{in} \quad G_4,$$

where $e_v = \pm 1$ and C_v are (possibly empty) reduced words on $a_1, \cdots, a_m$; suppose further that n is minimal (as we range over all words W of this form with $W = 1$ in G_4). Now $G_4 = G' * R$ and W involves only letters in G', so that

$$W = 1 \quad \text{in} \quad G'.$$

Since G' is a Britton extension of G_3' with stable letter k_0, Britton's lemma says that if W involves k_0, i.e., if $W \not\equiv C_0$, then W contains a pinch $k_0^e C_v k_0^{-e}$, where the group element corresponding to C_v is a word on $hr_i h^{-1}$, hxh^{-1}, and $hq^{-1}h^{-1}q_1 t_0 q_1^{-1} hqh^{-1}$. The relations of G' say that k_0 commutes with these words. Therefore, we may replace $k_0^e C_v k_0^{-e}$ in W by C_v, violating the minimality of n. It follows that W does not involve k_0, and therefore is a reduced word C_0 on $a_1, \cdots, a_m$. We have already seen that $\langle a_1, \cdots, a_m \rangle_4$ is free; hence $W \equiv 1$. ∎

Lemma 12.24 *G_5 is a Britton extension of G_4.*

Proof Clearly G_5 has basis G_4 and stable letters $b_1, \cdots, b_m$. To verify the isomorphism condition, we must show there are isomorphisms $\varphi_i\colon A(i) \to B(i)$, where

$$A(i) = \langle u_1, \cdots, u_m, a_1, \cdots, a_m, k_0 \rangle_4$$

and

$$B(i) = \langle u_1, \cdots, u_m, a_1, \cdots, a_m, k_0 u_i^{-1} \rangle_4,$$

and for which $\varphi_i(u_j) = u_j$, $\varphi_i(a_j) = a_j$, and $\varphi_i(k_0) = k_0 u_i^{-1}$. Note that the subgroups $A(i)$ and $B(i)$ are equal.

Visibly

$$\begin{aligned} A(i) &= \langle u_1, \cdots, u_m \rangle * \langle a_1, \cdots, a_m, k_0 \rangle_4 \\ &= R * \langle a_1, \cdots, a_m, k_0 \rangle_4, \end{aligned}$$

for $G_4 = R * G'$. Since $\langle a_1, \cdots, a_m, k_0 \rangle_4$ is free on the displayed generators, the map $\varphi_i\colon A(i) \to B(i) = A(i)$ described above is a well-defined homomorphism. Similarly, the map $\psi_i\colon A(i) \to A(i)$ given by

$$\psi_i(u_j) = u_j, \qquad \psi_i(a_j) = a_j, \qquad \text{and} \qquad \psi_i(k_0) = k_0 u_i$$

is also well defined. Since ψ_i is the inverse of φ_i, we have φ_i an isomorphism. ∎

Lemma 12.25 *G_6 is a Britton extension of G_5.*

Proof Clearly G_6 has basis G_5 and stable letter d. To verify the isomorphism condition, we must show that there is an isomorphism $\varphi\colon \langle k_0, a_1 b_1, \cdots, a_m b_m \rangle_5 \to \langle k_0, a_1, \cdots, a_m \rangle_5$ with $\varphi(k_0) = k_0$ and $\varphi(a_i b_i) = a_i$.

In G_5, $k_0^{-1} b_i k_0 = b_i u_i$. Thus, the function $G_5 \to G_4$, defined by setting each $b_i = 1$ and each $u_i = 1$ and sending the other generators to themselves, is a well defined homomorphism, for all the relations of G_5 are preserved ($b_i = 1$ implies $1 = k_0^{-1} b_i k_0 = b_i u_i = u_i$). This map takes each of $\langle k_0, a_1 b_1, \cdots, a_m b_m \rangle_5$ and $\langle k_0, a_1, \cdots, a_m \rangle_5$ onto the subgroup $\langle k_0, a_1, \cdots, a_m \rangle_4$ of G_4, which is free on the displayed generators (Lemma 12.23). By Exercise 12.16, each of the two subgroups of G_5 is free on the displayed generators. Therefore $k_0 \to k_0$ and $a_i b_i \to a_i$ defines an isomorphism. ∎

The next lemma will be needed in verifying that G_7 is a Britton extension of G_6.

Lemma 12.26 *The subgroup A of G' generated by $k_0, t_0, a_1, \cdots, a_m$ has the presentation*

$$A = (k_0, t_0, a_1, \cdots, a_m \mid k_0^{-1} w^{-1} t_0 w k_0 = w^{-1} t_0 w, \quad w \in E).$$

REMARK Recall our change in notation: Although E originally was a set of positive words on $u_1, \cdots, u_m$, we are now assuming E is comprised of positive words on $a_1, \cdots, a_m$.

Proof The relations $k_0^{-1} w^{-1} t_0 w k_0 = w^{-1} t_0 w$, for all $w \in E$, do hold in G', by Lemma 12.22, and hence they hold in the subgroup A of G'. To see that no other relations are needed, we shall show if U is a word on $k_0, t_0, a_1, \cdots, a_m$ for which

$$U = 1 \quad \text{in} \quad A,$$

then U can be transformed into 1 via elementary operations using only these relations.

It is easy to see that for every $w \in E$ and $\varepsilon = \pm 1$, $\eta = \pm 1$, the given relations imply

$$t_0^{\varepsilon} w k_0^{\eta} = w k_0^{\eta} w^{-1} t_0^{\varepsilon} w.$$

If U contains a subword of the form $t_0^{\varepsilon} w k_0^{\eta}$, then

$$U \equiv U_1 t_0^{\varepsilon} w k_0^{\eta} U_2 \to U_1 w k_0^{\eta} w^{-1} t_0^{\varepsilon} w U_2$$

is an elementary operation. Now cancel all subwords (if any) of the form yy^{-1} or $y^{-1}y$, where $y \equiv k_0$, t_0, or some a_i. With each such operation, the total number of occurrences of t_0^{ε} that precede k_0^{η}'s goes down. We may therefore assume that U contains no subword of the form $t_0^{\varepsilon} w k_0^{\eta}$, where $w \in E$, and that U is freely reduced.

If U does not involve k_0, then U is a reduced word on $t_0, a_1, \cdots, a_m$. But, by Lemma 12.23, the subgroup $\langle t_0, a_1, \cdots, a_m \rangle_4$ is free on these generators. Hence $U \equiv 1$, and the desired conclusion holds. Similarly, we are done if U does not involve t_0, for $\langle k_0, a_1, \cdots, a_m \rangle_4$ is free on these generators. Therefore, we may assume U involves both k_0 and t_0.

We are in a situation calling for Britton's lemma, for $U = 1$ in G', U involves k_0, and G' is a Britton extension of G_3' with stable letter k_0 (Lemma 12.21). Therefore, U contains a pinch $k_0^{\varepsilon} V k_0^{-\varepsilon}$, where

$$V = D \quad \text{in} \quad G_3'$$

and D is a word on hr_ih^{-1}, hxh^{-1}, and $hq^{-1}h^{-1}q_1t_0q_1^{-1}hqh^{-1}$. We assume that D is chosen with the minimal number of occurrences of t_0.

Recall that G_3' is a Britton extension of G_2 with stable letter t_0. We claim that both D and V are t_0-reduced. For notation, define $\Delta \equiv hq^{-1}h^{-1}q_1$, so that D is a word on hr_ih^{-1}, hxh^{-1}, and $\Delta t_0\Delta^{-1}$. If D is not t_0-reduced, it contains a pinch, and so

$$D \equiv D_1\Delta t_0^f\Delta^{-1}D_2\Delta t_0^{-f}\Delta^{-1}D_3,$$

where D_2 does not involve t_0. By Britton's lemma, we have

$$\Delta^{-1}D_2\Delta = N \quad \text{in} \quad G_2,$$

where N is a word on $q_1^{-1}hr_ih^{-1}q_1$ and $q_1^{-1}hxh^{-1}q_1$ (look at the presentation of G_3'); hence N commutes with t_0 in G_3'. Therefore, in G_3',

$$D = D_1\Delta t_0^fNt_0^{-f}\Delta^{-1}D_3 = D_1\Delta N\Delta^{-1}D_3,$$

contradicting our choice of D having the minimum number of occurrences of t_0. It follows that D is t_0-reduced.

We now show that V is t_0-reduced. If not, V contains a pinch $t_0^gCt_0^{-g}$ as a subword, where

$$C = N \quad \text{in} \quad G_2$$

for some word N (as in the paragraph above) that commutes with t_0 in G_3'. Now C is a subword of V not involving t_0, while V is a subword of U not involving k_0. Since U is a word on k_0, t_0, $a_1, \cdots, a_m$, we see that C is a word on $a_1, \cdots, a_m$. Since $G_3' \leq G_4$ and $\langle t_0, a_1, \cdots, a_m\rangle_4$ is free on the displayed generators, by Lemma 12.23, the element C commutes with t_0 in G_3' (hence in G_4) if and only if $C \equiv 1$. The pinch $t_0^gCt_0^{-g}$ in V is thus $t_0^gt_0^{-g}$; since V is a subword of U, we have contradicted the assumption that U is freely reduced.

We claim next that D must involve t_0. As $V = D$ in G_3' and both V and D are t_0-reduced, Corollary 12.9 applies to show that V and D have the same number of occurrences of t_0. If we assume that D does not involve t_0, then V does not involve t_0. Therefore, V is a word on $a_1, \cdots, a_m$, and D is a word on hr_1h^{-1} and hxh^{-1} (we assume D is chosen with the minimal total number of occurrences of all r_i). The equation $V = D$ thus holds in G_2, which is a Britton extension of $G_1 * \langle q, q_0, \cdots, q_N\rangle$ with stable letters r_i. Visibly V

is r_i-reduced, all i, for it is a word on $a_1, \cdots, a_m$, and hence does not even involve any r_i. To see that D is r_i-reduced for all i, assume otherwise. As D is a word on hr_ih^{-1} and hxh^{-1} containing a pinch,

$$D \equiv D_1 hr_i^l h^{-1} D_2 hr_i^{-l} h^{-1} D_3,$$

where D_2 involves no r_i's. Thus, $D_2 = hx^m h^{-1}$, and we have

$$D = D_1 hr_i^l h^{-1} x^m r_i^{-l} hD_3 \quad \text{in} \quad G_2.$$

The pinch in D is thus $r_i^l x^m r_i^{-l}$, and Britton's lemma asserts that x^m is equal in $G_1 * \langle q, q_0, \cdots, q_N \rangle$ to a word on $\bar{F}_i q_{i_1} G_i, s_1 x, \cdots, s_M x$ or a word on $\bar{H}_i q_{i_2} K_i, s_1 x^{-1}, \cdots, s_M x^{-1}$. But neither possibility can occur unless $m = 0$, as we have seen in the proof of Lemma 12.17. Therefore, we have $m = 0$ and we may erase $r_i^l r_i^{-l}$ from D, contradicting our assumption that the total number of occurrences of all r_i in D is minimal. We conclude that both V and D are r_i-reduced for all i, so that Corollary 12.9 says that V and D have the same number of occurrences of each r_i. Therefore D involves no r_i and is thus a word on hxh^{-1}, i.e., we have $D = hx^n h^{-1}$. The equation $V = D$ thus holds in G_1 and is

$$V = hx^n h^{-1} \quad \text{in} \quad G_1.$$

Suppose V involves a_j. Since G_1 is a Britton extension of $\langle x \rangle$ with stable letters $h, s_0, \cdots, s_M$, and since $\{a_1, \cdots, a_m\} \subset \{s_0, \cdots, s_M\}$, Britton's lemma asserts that $Vhx^{-n}h^{-1}$ contains a pinch $a_j^d C a_j^{-d}$, where C is a word on x. Since $h \notin \{s_0, \cdots, s_M\}$, h is not a_j, so that this pinch must be a subword of V. But V does not involve x, so that $C \equiv 1$, contradicting the fact that V is freely reduced. As V is a word on $a_1, \cdots, a_m$, we must have $V \equiv 1$. This is another contradiction, for $k_0^e V k_0^{-e} \equiv k_0^e k_0^{-e}$ is a subword of U, and U is freely reduced.

We have proved that D involves t_0, so that D involves $hq^{-1}h^{-1}q_1 t_0 q_1^{-1} hqh^{-1}$. Let us write

$$D \equiv D_1 (hq^{-1}h^{-1}q_1 t_0^{\alpha} q_1^{-1} hqh^{-1}) T,$$

where $\alpha = \pm 1$, the word in parentheses is the final occurrence of the long conjugate of t_0 occurring in D, and T is a word on hr_ih^{-1} and hxh^{-1}. Since D involves t_0, we know that V also involves t_0, so we may write

$$V \equiv V_0 t_0^{\varepsilon_1} \cdots V_{r-1} t_0^{\varepsilon_r} V_r,$$

where each V_j is a word on $a_1, \cdots, a_m$. We noted earlier that V and D are each t_0-reduced, so we may apply Corollary 12.9 to the equation $V = D$ in G_3'. It follows that

$$t_0^{\varepsilon_r} V_r T^{-1} h q^{-1} h^{-1} q_1 t_0^{-\alpha}$$

is a pinch, so that

$$V_r T^{-1} h q^{-1} h^{-1} q_1 = K \quad \text{in} \quad G_2,$$

where K is a word on $q_1^{-1} h r_i h^{-1} q_1$ and $q_1^{-1} h x h^{-1} q_1$. Since T is a word on $h r_i h^{-1}$ and $h x h^{-1}$, we may write

$$T^{-1} = h L_1 h^{-1} \quad \text{in} \quad G_2,$$

where L_1 is a word on r_i and x. We may write

$$K = q_1^{-1} h L_2 h^{-1} q_1 \quad \text{in} \quad G_2,$$

where L_2 is also a word on r_i and x. Therefore, we have

$$V_r h L_1 h^{-1} h q^{-1} h^{-1} q_1 = q_1^{-1} h L_2 h^{-1} q_1 \quad \text{in} \quad G_2,$$

which we may rewrite as

$$L_2^{-1} h^{-1} q_1 V_r h L_1 = q \quad \text{in} \quad G_2.$$

Note that $V_r h$ is freely reduced, for V_r is a freely reduced word on $\{a_1, \cdots, a_m\}$ and $h \notin \{a_1, \cdots, a_m\}$. Lemma 12.18 applies to give $V_r h$, hence its subword V_r positive, and

$$h q_1 V_r h = q \quad \text{in} \quad \gamma.$$

Since V_r is a positive word on S, Lemma 12.3 says that $V_r \in E$. Let us return to the birthplace of V_r: the word $t_0^{\varepsilon_r} V_r$ is a subword of V, and $k_0^{e} V k_0^{-e}$ is a subword of U. Thus $t_0^{\varepsilon_r} V_r k_0^{-e}$ is a subword of U, and, since $V_r \in E$, we have contradicted the condition that U contains no such subword. ∎

Lemma 12.27 *G_7 is a Britton extension of G_6.*

Proof Clearly G_7 has basis G_6 and stable letter σ. To verify the isomorphism condition, we must show there is an isomorphism

$$\varphi\colon \langle k_0, t_0, a_1, \cdots, a_m \rangle_6 \to \langle k_0, t_0 d, a_1, \cdots, a_m \rangle_6$$

with $\varphi(k_0) = k_0$, $\varphi(t_0) = t_0 d$, and $\varphi(a_i) = a_i$. Since $G' \leq G_5 \leq G_6$, we have $\langle k_0, t_0, a_1, \cdots, a_m \rangle_6 = \langle k_0, t_0, a_1, \cdots, a_m \rangle_4 = A$,

the subgroup whose presentation was just determined in Lemma 12.26. Let us denote the subgroup $\langle k_0, t_0 d, a_1, \cdots, a_m \rangle_6$ by B. To show that the map $\varphi\colon A \to B$ is a homomorphism, we must check whether it preserves the relations of A: If $w \in E$, then is

$$k_0^{-1} w^{-1} t_0 d w k_0 = w^{-1} t_0 d w \quad \text{in} \quad B?$$

We shall show this equation holds in G_6 (which implies that it holds in B).

Let us introduce notation. If w is a word on $a_1, \cdots, a_m$, then w_b is the word obtained from w by replacing each a_i by b_i; w_u is the word obtained from w by replacing each a_i by u_i. If $w \in E$, then $w_u = 1$, for w_u is one of the original defining relations of R. For $w \in E$, each of the following equations holds in G_6.

$$\begin{aligned} k_0^{-1} w^{-1} t_0 d w k_0 &= k_0^{-1} w^{-1} t_0 (d w d^{-1}) d k_0 \\ &= k_0^{-1} w^{-1} t_0 w w_b d k_0 \end{aligned}$$

(because $d a_i d^{-1} = a_i b_i$ in G_6 and a_i, b_j commute in $G_5 \leq G_6$). Since k_0 and d commute in G_6, we have

$$\begin{aligned} k_0^{-1} w^{-1} t_0 w w_b d k_0 &= k_0^{-1} w^{-1} t_0 w w_b k_0 d \\ &= k_0^{-1} w^{-1} t_0 w k_0 (k_0^{-1} w_b k_0) d \\ &= k_0^{-1} w^{-1} t_0 w k_0 w_b w_u d \end{aligned}$$

(because b_i and u_j commute and $k_0^{-1} b_i k_0 = b_i u_i$)

$$= k_0^{-1} w^{-1} t_0 w k_0 w_b d$$

(since $w_u = 1$). We have shown that

$$\begin{aligned} k_0^{-1} w^{-1} t_0 d w k_0 &= (k_0^{-1} w^{-1} t_0 w k_0) w_b d \\ &= w^{-1} t_0 w w_b d. \end{aligned}$$

On the other hand,

$$\begin{aligned} w^{-1} t_0 d w &= w^{-1} t_0 (d w d^{-1}) d \\ &= w^{-1} t_0 w w_b d, \end{aligned}$$

as we saw above. Therefore

$$k_0^{-1} w^{-1} t_0 d w k_0 = w^{-1} t_0 d w \quad \text{in} \quad G_6$$

and $\varphi\colon A \to B$ is a well-defined homomorphism onto B.

To see that φ is an isomorphism, we construct a homomorphism $\psi\colon G_6 \to G_6$ whose restriction $\psi|B$ is the inverse of φ. Define ψ by setting

$$\psi(d) = \psi(b_i) = \psi(u_i) = 1$$

and $\psi|G' = 1_{G'}$. Inspection of the various presentations shows that $\psi\colon G_6 \to G_6$ is a well defined homomorphism. Since $\psi(k_0) = k_0$, $\psi(a_i) = a_i$, and $\psi(t_0 d) = t_0$, we see that $\psi|B$ is the inverse of φ. ∎

We have now verified that $R \subset G_4 \leq G_5 \leq G_6 \leq G_7$, so that R is imbedded in G_7. Note that G_7 is finitely generated, but the presentation we have been working with has infinitely many relations.

Lemma 12.28 *G_7 is finitely presented.*

Proof If we delete those relations in G_7 of the form $w_u = 1$, where $w \in E$, only a finite number of relations remain (recall that w_u means replace each a_i in w by u_i); we claim that $w_u = 1$ is a consequence of remaining relations.

If $w \in E$, then $k_0^{-1} w^{-1} t_0 w k_0 = w^{-1} t_0 w$. Therefore,

$$\sigma^{-1}(k_0^{-1} w^{-1} t_0 w k_0)\sigma = \sigma^{-1} w^{-1} t_0 w \sigma.$$

Since σ commutes with k_0 and all a_i, this gives

$$k_0^{-1} w^{-1} \sigma^{-1} t_0 \sigma w k_0 = w^{-1} \sigma^{-1} t_0 \sigma w.$$

Therefore

$$k_0 w^{-1} t_0 d w k_0 = w^{-1} t_0 d w.$$

Inserting $w k_0 k_0^{-1} w^{-1}$ and $w w^{-1}$ gives

$$(k_0 w^{-1} t_0 w k_0) k_0^{-1} w^{-1} d w k_0 = (w^{-1} t_0 w) w^{-1} d w$$

The terms in parentheses are equal, so that cancelling gives

$$k_0^{-1} w^{-1} d w k_0 = w^{-1} d w. \tag{5}$$

Now the relations $d^{-1} a_i b_i d = a_i$ and $a_i b_j = b_j a_i$ give

$$d w d^{-1} = w w_b$$

or

$$w^{-1} d w = w^{-1} d d^{-1} w w_b d = w_b d. \tag{6}$$

Substituting equation (6) into equation (5) gives $k_0^{-1}w_b dk_0 = w_b d$. Since k_0 and d commute, cancellation of d gives

$$k_0^{-1}w_b k_0 = w_b. \tag{7}$$

But the relations $k_0^{-1}b_i k_0 = b_i u_i$ and $b_i u_j = u_j b_i$ give

$$k_0^{-1}w_b k_0 = w_b w_u.$$

This last equation, together with (7), gives

$$w_b = w_b w_u,$$

so that $w_u = 1$, as desired. ∎

We have thus imbedded the given recursively presented group R into G_7, a finitely presented group. This completes the proof of G. Higman's imbedding theorem. ∎

The next corollary may allow us to drop the hypothesis that R be finitely generated. If G is a countable group, let G^{II} denote the group on 2 generators containing G that was constructed in Theorem 11.37.

Corollary 12.29 *If G is a countable group for which G^{II} is recursively presented, then G can be imbedded in a finitely presented group.*

Proof Higman's theorem asserts that G^{II} can be imbedded in a finitely presented group. ∎

At this point, we omit some details which essentially require accurate bookkeeping in order to give a presentation for G^{II} from a given presentation of G. We assert there is a presentation of the abelian group

$$G = \sum_{\aleph_0} (\mathbf{Q} \oplus \mathbf{Q}/Z)$$

for which G^{II} is recursively presented.

Corollary 12.30 *There exists a finitely presented group that contains an isomorphic copy of every countable abelian group as a subgroup.*

Proof By Exercise 9.46, every countable abelian group may be imbedded in

$$G = \sum_{\aleph_0} (\mathbf{Q} \oplus \mathbf{Q}/Z).$$

Our assertion above is that G^{II} is recursively presented, so that the result follows from Corollary 12.29. ▮

There are only countably many finitely presented groups, and their free product is a countable group G having a presentation for which G^{II} is recursively presented.

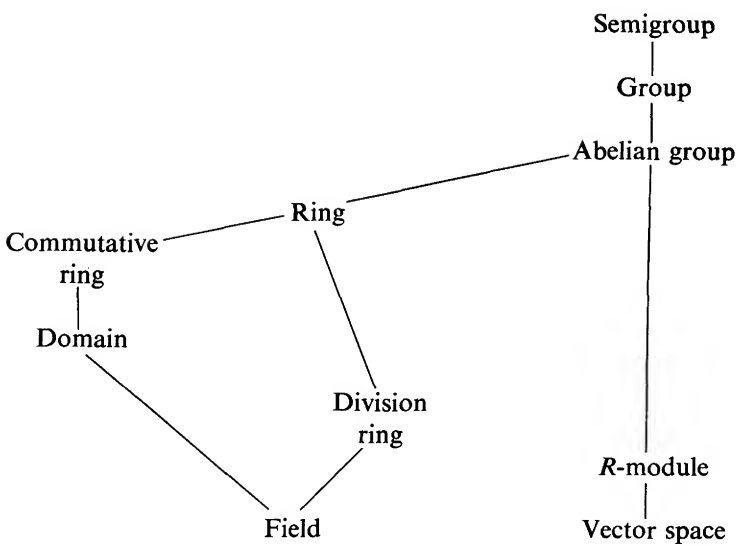
Corollary 12.31 *There exists a finitely presented group containing an isomorphic copy of every finitely presented group as a subgroup.*

Proof The result follows from Corollary 12.29 and the assertion above about the free product G. ▮

Let us now indicate how the Boone-Novikov theorem may be derived from the Higman theorem. It is not difficult to construct a recursively presented group G having an unsolvable word problem (there is such a construction that is a variant of the group exhibited in Theorem 12.11). By Higman's theorem, there is a finitely presented group H containing G. Since every subgroup of a group having a solvable word problem also has a solvable word problem, it follows that the word problem for H is unsolvable.

APPENDIX
I

Some Major Algebraic Systems



A *ring* (which we always assume to contain a unit $1 \neq 0$) is a set with two binary operations: addition and multiplication. It is an abelian group under addition, a semigroup under multiplication, and the two operations are linked by the distributive laws.

A *commutative ring* is a ring in which multiplication is commutative.

A *domain* (or integral domain) is a commutative ring having no zero divisors, i.e., nonzero elements a, b with $ab = 0$.

A *field* is a commutative ring in which every nonzero element has a multiplicative inverse.

A *division ring* (or skew field) is a ring in which every nonzero element has a multiplicative inverse. Thus, a commutative division ring is a field. The nonzero elements of a division ring form a multiplicative group.

If R is a ring, an abelian group M is an *R-module* if there is a scalar multiplication defined, i.e., if there is a function $R \times M \to M$ (whose value on (r, m) is denoted rm), which has the properties:

$$(rr')m = r(r'm);$$

$$r(m + m') = rm + rm';$$

$$(r + r')m = rm + r'm;$$

$$1m = m$$

for all r, r', $1 \in R$ and m, $m' \in M$.

A *vector space* is an R-module where R is a field.

APPENDIX

II

Equivalence Relations and Equivalence Classes

If X is a nonempty set, a *binary relation* on X is a subset R of $X \times X$. If $(x, y) \in X \times X$, one writes xRy instead of $(x, y) \in R$. For example, the binary relation $<$ on the reals consists of all points in the plane lying above the line $y = x$. Ordinarily, one writes $2 < 3$ instead of $(2, 3) \in <$.

A binary relation $\sim$ on X is an *equivalence relation* in case, for all $x, y, z \in X$:

(i) $x \sim x$;
(ii) $x \sim y$ implies $y \sim x$;
(iii) $x \sim y$ and $y \sim z$ implies $x \sim z$.

If $x \in X$, we let $[x]$ denote the family of all $y \in X$ such that $y \sim x$; we call $[x]$ the *equivalence class* containing x.

Let X be a nonempty set. A *partition* of X is a family of nonempty subsets $\{S_i : i \in I\}$ of X such that:

$$S_i \cap S_j = \varnothing \qquad \text{if } i \neq j \text{ (\textit{pairwise disjointness})};$$

$$X = \bigcup S_i.$$

Proposition A *If $\sim$ is an equivalence relation on X, then the family of all equivalence classes is a partition of X.*

Proof By (i), if $x \in X$, then $x \in [x]$; thus, the equivalence classes are nonempty. Moreover,

$$X = \bigcup_{x \in X} [x],$$

since every $x \in X$ lies in its own equivalence class.

Suppose $[x] \cap [y] \neq \varnothing$, so that there is an element $z \in [x] \cap [y]$. Then $z \sim x$ and $z \sim y$. By (ii), $x \sim z$; therefore, $x \sim y$, by (iii), and $[x] = [y]$. ∎

Proposition B *If $\{S_i : i \in I\}$ is a partition of X, then there is an equivalence relation on X whose equivalence classes are the S_i.*

Proof Define $x \sim y$ if there is an S_i containing x and y. It is immediate that (i) and (ii) hold. To prove (iii), suppose $x, y \in S_i$ and $y, z \in S_j$. Then $y \in S_i \cap S_j$, so that $S_i = S_j$ (pairwise disjointness) and $x \sim z$. Therefore we have defined an equivalence relation. If $x \in S_i$, it follows easily that $[x] = S_i$. ∎

The importance of equivalence relations is just this: Let $\sim$ be an equivalence relation on X and let Y be the set of equivalence classes. If one treats the elements of Y merely as elements (and not as classes), then he has effectively identified equivalent elements of X.

For example, one does not wish to distinguish between the fractions $\frac{1}{2}$ and $\frac{2}{4}$, and so he decrees that two fractions a/b and c/d are "equal" if $ad = bc$. In reality, he is considering the set X of all ordered pairs of integers (a, b) for which $b \neq 0$, under the equivalence relation given by

$$(a, b) \sim (c, d) \qquad \text{if } ad = bc.$$

The rational $\frac{1}{2}$ is thus the class of (1, 2) and of (2, 4).

APPENDIX III

Functions

If X and Y are sets, a *relation* from X to Y is a subset of $X \times Y$. A *function* $f\colon X \to Y$ is a relation from X to Y such that:

(i) For each $x \in X$, there is some $y \in Y$ with $(x, y) \in f$;
(ii) For each $x \in X$, the y above is unique.

If $(x, y) \in f$, the element y is the *image* of x under f and is ordinarily written $f(x)$. With this notation, a function f is the subset of $X \times Y$ consisting of all pairs $(x, f(x))$. In other words, a function is defined by its domain X, its target Y, and its graph.

In practice, one thinks of a function as something dynamic; it assigns elements of Y to elements of X. Indeed, most elementary texts define a function as a "rule of correspondence". Even at a naive level, though, this can be misleading. For example, are $(x + 1)^2$ and $x^2 + 2x + 1$ different rules? The reader may prove, using the definition above, that if f and $g\colon X \to Y$, then $f = g$ if and only if $f(x) = g(x)$ for each $x \in X$.

Part (ii) of the definition of function deserves a bit more comment; it says that a function is "single-valued", or, as we prefer to say, that a function is *well defined.* When attempting to define a function, one must prove that (ii) holds, lest he define only a relation.

The function $f\colon X \to X$ defined by $f(x) = x$ for all $x \in X$ is called the *identity function* on X, and f is denoted 1_X. If Y is a subset of X, the function $i\colon Y \to X$ defined by $i(y) = y$ for all $y \in Y$ is called the *inclusion.*

If Y is a proper subset of X, then 1_Y is distinct from the inclusion i, for $1_Y: Y \to Y$ while $i: Y \to X$. Thus, in order that two functions be the same, it is not enough that they have the same graph; they must also have the same domains and same target sets.

If $i: Y \to X$ is an inclusion and $f: X \to Z$ is a function, then the *restriction* of f to Y, denoted $f|Y$, is the composite $f \circ i$. Thus, $f|Y: Y \to Z$, and, for all $y \in Y$, we have $(f|Y)(y) = f(y)$.

A function $f: X \to Y$ is *one-to-one* if distinct elements of X have distinct images, i.e., if $f(x) = f(x')$, then $x = x'$. This definition may profitably be compared with that of f being well defined, which is the converse: If $x = x'$, then $f(x) = f(x')$. A function $f: X \to Y$ is *onto* if each $y \in Y$ is the image of something in X; i.e., if $y \in Y$, there exists an $x \in X$ with $f(x) = y$. A function $f: X \to Y$ is a *one-to-one correspondence* if it is both one-to-one and onto; such a function always has an *inverse*, i.e., a function $g: Y \to X$ such that $f \circ g = 1_Y$ and $g \circ f = 1_X$. One must consider both composites, for $f \circ g = 1_Y$ implies only that f is onto and g is one-to-one.

Two sets X and Y (possibly infinite) have the *same number of elements* if there is a one-to-one correspondence $f: X \to Y$. Let N be the set of positive integers; a set X is *countable* if it is finite or if it has the same number of elements as N. An infinite set X is countable, therefore, if and only if there is a list of its elements $x_1, x_2, \cdots, x_n, \cdots$ (define $x_n = f(n)$, where $f: N \to X$ is a one-to-one correspondence). The set Z of all integers (positive, negative, and zero) is countable as is the set $\mathbf{Q}$ of all rational numbers. The set $\mathbf{R}$ of real numbers and the set $\mathbf{C}$ of complex numbers are uncountable.

If X and Y are sets, we write $|X| \leq |Y|$ in case X has the same number of elements as a subset of Y; i.e., there is a function $f: X \to Y$ that is one-to-one (but not necessarily onto). The *Cantor-Bernstein theorem* states that if $|X| \leq |Y|$ and $|Y| \leq |X|$, then $|X| = |Y|$; i.e., X and Y have the same number of elements.

APPENDIX IV

Zorn's Lemma

Let X be a nonempty set. A binary relation $\leq$ on X is a *partial order* in case, for $x, y, z \in X$:

$$\begin{aligned} &x \leq y; \\ &x \leq y \quad \text{and} \quad y \leq x \quad \text{implies} \quad x = y; \\ &x \leq y \quad \text{and} \quad y \leq z \quad \text{implies} \quad x \leq z. \end{aligned}$$

The best example of a partially ordered set is a collection of subsets of a set Y, where $\leq$ means $\subset$.

A partial order is a *simple order* (or total order) if, for each $x, y \in X$, either

$$x \leq y \qquad \text{or} \qquad y \leq x.$$

If X is a partially ordered set, a *chain* is a simply ordered subset. For example, the rational numbers form a chain in the real numbers.

If S is a nonempty subset of X, an *upper bound* of S is an element $x_0 \in X$ (not necessarily in S) such that

$$s \leq x_0 \qquad \text{for all } s \in S.$$

Finally, a *maximal element* in X is an element y_0 which is smaller than no other element in X; i.e.,

$$\text{if } y_0 \leq x, \qquad \text{then } y_0 = x.$$

There are partially ordered sets having many maximal elements, and there are partially ordered sets having no maximal elements.

Zorn's Lemma *Let X be a partially ordered set in which every chain has an upper bound. Then there exists a maximal element in X.*

Zorn's lemma is equivalent to a much more intuitive statement, the axiom of choice, which says that the cartesian product of nonempty sets is itself nonempty. We regard either of these statements as an axiom of mathematics, and we shall not be ashamed to use either when necessary. There is another statement equivalent to the axiom of choice which is often useful. A partially ordered set X is called *well-ordered* if every nonempty subset of X contains a smallest element. (Well-ordered sets must be simply ordered.) The positive integers N is well-ordered, but the set Z of all integers is not well-ordered.

Well-ordering Principle. *Given a nonempty set X, there is a partial order on X for which X is well-ordered.*

For example, though Z is not well-ordered under the usual definition of $\leq$, a new ordering can be defined on it so that it is well-ordered $\{0, 1, -1, 2, -2, \cdots\}$. Here is an example of a well-ordered set in which an element may have infinitely many predecessors: Let X be the following subset of $\mathbf{R}$ with the usual definition of $\leq$:

$$X = \left\{1 - \frac{1}{n} : n > 0\right\} \cup \left\{2 - \frac{1}{n} : n > 0\right\}.$$

APPENDIX V

Principal Ideal Domains

We shall let R denote a commutative ring with unit throughout this appendix.

If $x_1, \cdots, x_k$ are elements of R, let $(x_1, \cdots, x_k)$ denote the set of all linear combinations of the x with coefficients in R:

$$(x_1, \cdots, x_k) = \{\sum r_i x_i : r_i \in R\}.$$

It is easy to check that $(x_1, \cdots, x_k)$ is an ideal in R.

If a and b are in R, one says that *a divides b* in case $ac = b$ for some $c \in R$. If $x_1, \cdots, x_k$ are in R, a *common divisor* of $x_1, \cdots, x_k$ is an element $c \in R$ that divides each x_i; a *greatest common divisor* (gcd) is a common divisor that is divisible by every common divisor.

A *principal ideal domain* (*PID*) is a domain in which every ideal is principal. For each ideal I in a *PID* R, therefore, there is an element $r_0 \in I$ with

$$I = (r_0) = \{rr_0 : r \in R\}.$$

Theorem A *If R is a PID and if $x_1, \cdots, x_k$ are elements of R, then R contains a gcd of $x_1, \cdots, x_k$, and this gcd is a linear combination of the x.*

Proof Since R is a *PID*, there is an element $d \in R$ with

$$(x_1, \cdots, x_k) = (d);$$

as any element of $(x_1, \cdots, x_k)$, d is a linear combination of the x.

It follows that any common divisor of the x divides d. But d is a common divisor of the x, for each $x_i \in (d)$, so that $x_i = r_i d$ for some $r_i \in R$. Therefore, d is a gcd. ∎

A *unit* in R is an element $u \in R$ that has a multiplicative inverse in R; i.e., there is an element $v \in R$ with $uv = 1$. Two elements a and b in R are *associates* if there is a unit $u \in R$ with $a = ub$.

Theorem B *Let R be a PID, and let $x_1, \cdots, x_k \in R$. Any two gcd of $x_1, \cdots, x_k$ are associates.*

Proof If a and b are gcd's, then each divides the other. Therefore, $a = ub$ and $b = va$, where $u, v \in R$. Hence, $a = uva$, so that $1 = uv$, since R is a domain. Therefore, u is a unit and a and b are associates. ∎

An element $p \in R$ is *irreducible* in case p is not a unit, and in every factorization $p = ab$, either a or b is a unit.

A domain R is a *unique factorization domain* (*UFD*) in case:

(i) Every nonzero $a \in R$ that is not a unit is a product of irreducible elements;

(ii) If $p_1 \cdots p_m = q_1 \cdots q_n$, where the p and q are irreducibles, then there is a one-to-one correspondence between the factors (i.e., $m = n$) such that corresponding factors are associates.

In short, R is a *UFD* if the fundamental theorem of arithmetic holds in R.

We wish to prove that if R is a *PID*, then R is a *UFD*; our first task is to show that every nonzero $a \in R$ that is not a unit is a product of irreducibles.

Lemma C *If R is a PID, there is no infinite sequence of ideals*

$$I_1 \subsetneq I_2 \subsetneq \cdots \subsetneq I_n \subsetneq I_{n+1} \subsetneq \cdots.$$

Proof It is easy to check that

$$I = \bigcup_{n=1}^{\infty} I_n$$

is an ideal. Since R is a *PID*, $I = (d)$ for some $d \in R$. Now d got into I by being in I_n for some n. Hence,

$$I = (d) \subset I_n \subsetneq I_{n+1} \subset I,$$

a contradiction. ∎

Lemma D *If R is a PID and $a \in R$ is nonzero and not a unit, then a is a product of irreducibles.*

Proof If $a = bc$, where neither b nor c is a unit, then we say that b is a *proper factor* of a. It is easy to check that if b is a proper factor of a, then $(a) \subsetneq (b)$.

Call a "good" if it is a product of irreducibles; otherwise, a is "bad". If b and c are good, so is their product bc. Thus, if a is bad, a is not irreducible and it has a proper bad factor. Suppose $a = a_0$ is bad. Assume inductively that there exist $a_0, a_1, \cdots, a_n$ such that each a_{i+1} is a bad, proper factor of a_i. Since a_n is bad, it has a proper, bad factor a_{n+1}. By induction, there is an infinite sequence $a_0, a_1, \cdots$, in which each a_{i+1} is a proper factor of a_i. There is thus an infinite sequence of ideals $(a_0) \subsetneq (a_1) \subsetneq \cdots$, and this is a contradiction. Therefore, every nonzero nonunit is good. ∎

Theorem E (Euclid) *Let R be a PID and let p be an irreducible element in R. If p divides ab, then p divides a or p divides b.*

Proof If p does not divide a, then the gcd of p and a is 1, for p is irreducible. By Theorem A, there are elements $s, t \in R$ with $1 = sa + tp$. Therefore, $b = sab + tpb$. Since p divides ab, we see that p divides each of the terms on the right; hence p divides b. ∎

Theorem F (Fundamental Theorem of Arithmetic) *Every principal ideal domain R is a unique factorization domain.*

Proof By Lemma D, every nonzero $a \in R$ that is not a unit is a product of irreducibles.

If $p_1 \cdots p_m = q_1 \cdots q_n$, where the p and q are irreducibles, then p_1 divides $q_1 \cdots q_n$. By iterated applications of Theorem E, p_1 divides some q_j. Since both p_1 and q_j are irreducibles, they must be associates: $p_1 u = q_j$ for some unit u. Since R is a domain,

$$(up_2)p_3 \cdots p_m = \prod_{i \neq j} q_i,$$

and the proof is completed by an induction on $\max\{m, n\}$. ∎

BIBLIOGRAPHY

Artin, E., *Geometric Algebra*, Interscience, 1957.

———, *Galois Theory*, Notre Dame, 1955.

Babakhanian, A., *Cohomological Methods in Group Theory*, M. Dekker, 1972.

Birkhoff, G., and Mac Lane, S., *A Survey of Modern Algebra*, Macmillan, 1941.

Burnside, W., *The Theory of Groups of Finite Order*, Cambridge University Press, 1911.

Carmichael, R., *An Introduction to the Theory of Groups of Finite Order*, Ginn, 1937.

Cohn, P. M., *Universal Algebra*, Harper & Row, 1965.

Coxeter, H. S. M., and Moser, W. O., *Generators and Relations for Discrete Groups*, Springer-Verlag, 1965.

Curtis, C., and Reiner, I., *Representation Theory of Finite Groups and Associative Algebras*, Wiley, 1962.

Davis, M., *Computability and Unsolvability*, McGraw-Hill, 1958.

Dickson, L. E., *Linear Groups*, Leipzig, 1900.

Dieudonné, J., *Sur les Groupes Classiques*, Hermann, 1958.

Dixon, J. D., *Problems in Group Theory*, Blaisdell, 1967.

Feit, W., *Characters of Finite Groups*, Benjamin, 1969.

Fuchs, L., *Infinite Abelian Groups*, Academic, 1970.

Gorenstein, D., *Finite Groups*, Harper & Row, 1968.

Griffith, P. A., *Infinite Abelian Group Theory*, University of Chicago, 1970.

Gruenberg, K. W., *Cohomological Topics in Group Theory*, Springer-Verlag, 1970.
Hall, M., Jr., *The Theory of Groups*, Macmillan, 1959.
Huppert, B., *Endliche Gruppen*, Springer-Verlag, 1967.
Jacobson, N., *Lectures in Abstract Algebra, I, II, III*, Van Nostrand, 1951.
Kaplansky, I., *Infinite Abelian Groups*, University of Michigan, 1953.
Kurosh, A. G., *The Theory of Groups*, Chelsea, 1956.
Ledermann, W., *Introduction to the Theory of Finite Groups*, Oliver and Boyd, 1957.
Lyndon, R., and Schupp, P., *Combinatorial Group Theory*, Ergebnisse Series, Springer-Verlag, to appear 1974.
Mac Lane, S., *Homology*, Academic, 1963.
Magnus, W., Karrass, A., and Solitar, D., *Combinatorial Group Theory*, Wiley, 1966.
Massey, W. S., *Algebraic Topology: An Introduction*, Harcourt, Brace and World, 1967.
Miller, C. F., III, *On Group-Theoretic Decision Problems and their Classification*, Princeton, 1971.
Miller, G. A., Blichfeldt, H. F., and Dickson, L. E., *Theory and Applications of Finite Groups*, Wiley, 1916.
Neumann, H., *Varieties of Groups*, Springer-Verlag, 1967.
Passman, D. S., *Permutation Groups*, Benjamin, 1969.
Robinson, D. J. S., *Finiteness Conditions and Generalized Soluble Groups*, Ergebnisse vols. 62 and 63, Springer-Verlag, 1972.
Rotman, J. J., *Notes on Homological Algebra*, van Nostrand-Reinhold, 1970.
Schenkman, E., *Group Theory*, van Nostrand, 1965.
Scott, W. R., *Group Theory*, Prentice-Hall, 1964.
Specht, W., *Gruppentheorie*, Springer-Verlag, 1956.
Van der Waerden, B. L., *Modern Algebra*, Ungar, 1948.
Wielandt, H., *Finite Permutation Groups*, Academic, 1964.
Zassenhaus, H., *The Theory of Groups*, Chelsea, 1956.

NOTATION

I. ALGEBRA

E	identity matrix
$GF(q)$	finite field with exactly q elements (q must be a prime power)
F^*	when F is a field, the multiplicative group of its nonzero elements
$R[x]$	when R is a ring, the ring of all polynomials in x with coefficients in R
Z_n	the integers modulo n

II. ELEMENTARY GROUP THEORY

$\{a_1, \cdots, a_n\}$	set with elements $a_1, \cdots, a_n$
$\langle a_1, \cdots, a_n \rangle$	subgroup generated by $a_1, \cdots, a_n$
ST	$\{st: s \in S \text{ and } t \in T\}$
$S \vee T$	subgroup generated by S and T
$[G:S]$	index of S in G
$\|G\|$	order of G
$H \lhd G$	H is a normal subgroup of G
$[a, b]$	$a^{-1}b^{-1}ab$
$[S, T]$	subgroup generated by all commutators $[s, t]$, where $s \in S$ and $t \in T$

III. SUBGROUPS

General Groups

G'	commutator subgroup of G
$Z(G)$	center of G
$C_G(x)$	centralizer of x in G
$C_G(H)$	centralizer of the subgroup H in G
$N_G(H)$	normalizer of subgroup H in G

Abelian Groups

G_p	p-primary component of G
nG	$\{nx: x \in G\}$
$G[n]$	$\{x \in G: nx = 0\}$
tG	torsion subgroup of G
dG	maximal divisible subgroup of G.

IV. NAMES OF SPECIAL GROUPS

A_n	alternating group of n letters
A_∞	infinite alternating group
$\mathbf{C}$	complex numbers
D_n	dihedral group of order $2n$
$GL(m, K)$	nonsingular $m \times m$ matrices over K
$GL(m, q)$	nonsingular $m \times m$ matrices over $GF(q)$
$SL(m, K)$	all $m \times m$ matrices over K with determinant 1
Z_0	center of $SL(m, K)$
$PSL(m, K)$	projective unimodular group $= SL/Z_0$
$\mathbf{Q}$	rational numbers
Q	quaternions
Q_n	generalized quaternions of order 2^n
$\mathbf{R}$	real numbers
S_n	symmetric group on n letters
S_X	all permutations of a set X
$\sigma(n)$	cyclic group of order n
$\sigma(p^\infty)$	all p^nth roots of unity, $n = 1, 2, \cdots$
$\mathbf{T}$	circle group
T	nonabelian group of order 12, not A_4, not D_6
V	Klein 4-group
Z	integers

Index

A_n = alternating group on n letters, 32
A_∞ = infinite alternating group, 41
Aanderaa, 308
Abelian group, 9
ACC (ascending chain condition), 76
Adjan-Novikov-Britton theorem, 66
Alphabet (of Turing machine), 281
Alternating group, 32
Ascending central series, 117
Automorphism
  of a field, 98
  of a group, 75
  group, 129
  inner, 129

Baer and Levi, 243
Baer sum, 228
Basic moves (of Turing machine), 281
Basic subgroup, 197
Basis
  free abelian group, 188
  free group, 238
  vector space, 180
Basis and stable letters, 291
Basis theorem
  finite, 54
  finitely generated, 193
Boone, 297
Boone's lemma, 299
Both chain conditions, 77
Bouquet of circles, 251
Britton extension, 293
Britton's lemma, 293
Burnside basis theorem, 126
Burnside $p^\alpha q^\beta$ theorem, 116

C = complex numbers
Cancellation law, 5
Cancellation semigroup, 9
Canonical decomposition
  group, 56
  module, 67
Canonical form
  Jordan, 71
  rational, 69
Carmichael, 38
Cauchy theorem, 83
Cayley theorem, 42
Cayley-Hamilton theorem:
  If $f(x)$ is the characteristic polynomial of a matrix A, then $f(A)$ is the zero matrix
Center, 34
  higher, 117
Centerless, 41
Central series, 120
Centralizes, 116
Centralizer
  of element, 35
  of subgroup, 116
Chain conditions, 76, 77
Character group, 216
Characteristic
  of element, 201
  of field, 153
Characteristic polynomial of A:
  $\det(xE - A)$
Characteristic root of A:
  a scalar c with $A\alpha = c\alpha$ for some nonzero vector α
Characteristic subgroup, 111
Characteristic vector of A:
  a nonzero vector α with $A\alpha = c\alpha$ for some scalar c

Church's thesis, 283
Circle group **T**, 8
Circuit, 247
Class equation, 84
Class of nilpotency, 118
Coboundary, 146, 218
Cocycle, 144n
Cohomology group, 146
Commute
  diagram, 174
  elements, 9
Commutator, 22
  subgroup, 22
    higher, 111
Companion matrix, 68
Complete group, 132
Complex, 243
  connected, 245
  covering, 252
Component
  complex, 245
  p-primary, 53
Composition factors, 107
Composition series, 104
Computation, 281
Conjugacy class, 34
Conjugate
  of element, 19
  of subgroup, 85
Conjugation, 8
Connected complex, 245
Connecting homomorphism, 230
Contravariant functor, 212
Correspondence theorem, 25
Coset, 14
  double, 17
Covariant functor, 211
Covering complex, 252
Crossed homomorphism, 149n
Cycle, 28
Cyclic group, 12
  module, 65

D_n = dihedral group of order $2n$, 89
DCC (descending chain condition), 77
Dedekind law, 23
Derivative, 155
Derived series, 111
Descending central series, 117
Dicyclic group Q_n = generalized quaternions, 94
Dihedral group D_n, 89
Dilatation, 165
Dimension
  complex, 244
  vector space, 182
Direct product
  finite, 49, 51
  infinite, 176
Direct sum
  finite, 52
  infinite, 177
  matrices, 69
  modules, 65
Direct summand, 53
Dirichlet, 151
Disjoint
  permutations, 29
  subcomplexes, 244
Divisible by n, 178
Divisible group, 179
Division algorithm:
  If a,b are integers with $a \neq 0$, there are unique integers q,r with $b = qa + r$ and $0 \leq r < |a|$
Domain (= integral domain), 61
Double coset, 17
Doubly transitive, 46
Dual diagram, 192
Dyadic rationals, 201

Edge, 244
Edgepath group, 247
Elementary abelian group, 57
Elementary divisors, 69
Elementary operations, 285
Empty word, 239
Endomorphism, 75
  nilpotent, 78
  normal, 76
Enumerates, 282
Equivalent
  characteristics, 202
  extensions, 146
  normal series, 104
  paths, 245
Euler φ-function, 17
Euler theorem, 18

Euler-Poincaré characteristic, 248n
Even permutation, 32
Exact sequence, 175
  short, 213
Exponent (of group), 57
Ext, 218
Extension, 127
  split, 136

Factor groups, 102
Factor set, 144, 217
Feit-Thompson theorem, 111
Fermat theorem, 18
Finitely generated module, 65
Finitely presented, 277
Fitting's lemma, 78, 82
Five lemma, 220
Fixes, 28, 98
4-group V, 38
Frattini argument, 88, 89
Frattini subgroup, 125
Free abelian group, 187
Free group, 238
Free product, 263
  with amalgamated subgroup, 270
Freely reduced word, 303
Full subcomplex, 244
Fully invariant subgroup, 111
Functional:
  a linear transformation $V \to K$, where V is a vector space over a field K
Functor
  contravariant, 212
  covariant, 211
  left exact, 213
Fundamental theorem
  arithmetic, 332
  finite abelian groups, 59
  finitely generated abelian groups, 193
  modules, 72

Galois field $GF(q)$, 156
Galois group, 99
gcd, 61, 330
General linear group $GL(m, K)$, 156
Generalized quaternions Q_n = dicyclic group, 94
Generators and relations, 236, 240
Golod-Šafarevič theorem, 66
Graph, 247
Grothendieck group, 205, 207
Group, 3
Group of units, 130

Hall subgroup, 115
P. Hall's theorem, 113
Hamel basis, 181
Hamiltonian group, 92
Height, 202n
G. Higman's theorem, 309
Higman-Neumann-Neumann theorems, 274, 275
HNN extension, 293n
Holomorph, 139
Hom, 212
Homomorphism
  group, 6
  module, 72
  crossed, 149n
Hyperplane, 165

Ideal, 60
  maximal, 63
  order, 66
  prime, 63
  principal, 60
Image, 12
Inclusion, 326
Indecomposable group, 75
Independence
  in group, 196
  in vector space, 180
Index, 15
Injective property, 183
Inner automorphism, 129
Instantaneous description, 280
Invariant subspace, 65
Isomorphism
  group, 6
  module, 72
Isomorphism condition, 292
Isomorphism theorem
  first, 21
  second, 24
  third, 24

Jordan block, 70
Jordan canonical form, 71
Jordan-Hölder theorem, 107

Kaplansky, 10, 200
Kernel, 12
Klein 4-group V, 38
Kulikov theorem, 197
Kuroš theorem, 267

Lagrange theorem, 16
Left exact functor, 213
Length
  cycle, 28
  normal series, 102
  path, 245
  word, 278
Lifting
  of element, 142
  of path, 253
Linear fractional transformation, 161

Magnus, 308
Mal'cev, 10
Mann, 18
Map (of complexes), 251
Markov, 279n
Maximal ideal, 63
Maximal normal subgroup, 26
Maximal tree, 247
Metabelian group, 118
Minimal normal subgroup, 112
Minimum polynomial, 69
Module, 64
Modular group, 265
Modular law, 23
Multiplication table, 6

Natural map, 22
Navel, Maurice (*see* Ombellico, Mario)
Neumann, 277
Nielsen-Schreier theorem, 258
Nilpotent endomorphism, 78
Nilpotent group, 118
  class of nilpotency, 118
Nongenerator, 126
Normal endomorphism, 76
Normal form, 271
Normal series, 102
Normal subgroup, 18
  generated by X, 19
  maximal, 26
  minimal, 112
Normalizer, 85
Normalizes, 116
Novikov-Boone theorem, 298
Nunke, 234

Obstruction, 232
Odd permutation, 32
Ombellico, Mario (*see* Pippik, Moishe)
Orbit, 30, 45
Order
  element, 16
  group, 15
Order ideal, 66

Parity, 34
Partial order, 328
Path, 245
  closed, 245
  reduced, 247
p-complement, 115
p-group, 53, 83
Penguin, 244
Permutation, 27
  disjoint, 29
  even, 32
  group = subgroup of S_X
  matrix, 33
  odd, 32
  regular, 29
Pinch, 294
Pippik, Moishe (*see* Navel, Maurice)
Poincaré, 44
Polya, 47
Positive word, 280
Post's theorem, 288
Presentation
  group, 240
  semigroup, 241
Primary component, 53
  decomposition, 53
  group, 53
Prime field, 152
Prime ideal, 63

Primitive element, 156
Principal ideal, 60
  domain, 61, 330
Product formula, 14
Projection, 75, 176, 253
Projective property, 189, 242
Projective space, 161
Projective unimodular group $PSL(m, K)$, 160
Proper subgroup, 12
Prüfer's theorem, 197
Pull-back, 223
Pure-independent subset, 196
Pure subgroup, 194
Push-out, 225

$\mathbf{Q}$ = rational numbers
Q_n = generalized quaternions of order 2^n, 94
Quadruple, 280
Quaternions Q of order 8, 91
Quotient complex, 252
Quotient group, 19
Quotient module, 65
Quotient ring, 62

$\mathbf{R}$ = real numbers
Rank
  free abelian group, 188
  free group, 243
  torsion-free abelian group, 200
Rational canonical form, 69
Realizes θ, 137, 143
Recursive, 283
Recursively enumerable (r.e.), 282
Recursively presented, 309
Reduced group, 185
  freely reduced word, 303
  reduced path, 247
  reduced word, 239
  p-reduced word, 295
Refinement, 103
Regular permutation, 29
  regular representation, 42
Remak-Krull-Schmidt theorem, 80
Representative (of coset), 14
Retract, 241
Retraction, 241
Root field, 97

S_X = all permutations on a set X, 27
S_n = symmetric group on n letters, 28
Schreier refinement theorem, 106
Schreier transversal, 260
Schur, 38
Schur-Zassenhaus lemma, 149
Semidirect product, 136
Semigroup, 3
  cancellation, 9
Series
  ascending central, 117
  central, 120
  composition, 104
  derived, 111
  descending central, 117
  normal, 102
  solvable, 108
Sheet, 253
Short exact sequence, 213
Similar:
  Two matrices A and B over a field F are similar if there is a nonsingular matrix P with $A = PBP^{-1}$. Two matrices A and B over a field F are similar if and only if they represent the same linear transformation T: $V \to V$ of a vector space V over F relative to (possibly) different ordered bases of V.
Simple group, 38
Simplex, 243, 244
Simultaneous bases, 193
Skeleton, 244
Solvable
  group, 102, 108
  by radicals, 98
  series, 108
  word problem, 278, 285
Special linear group $SL(m, K)$, 157
Special word, 286, 298
Split extension, 136
Stabilizer, 45
Stable letters, 291
Stopping state, 287
Subcomplex, 244
  full, 244

Subgroup, 11
  basic, 197
  characteristic, 111
  commutator, 22
  Frattini, 125
  fully invariant, 111
  generated by X, 13
  Hall, 115
  normal, 18
  proper, 12
  pure, 194
  subnormal, 104
  sylow, 85
  torsion, 175
Submodule, 65
Subnormal subgroup, 104
Sylow subgroup, 85
Sylow theorems, 87, 88
Symmetric group S_n, 28
$\sigma(n)$ = cyclic group of order n, 50

$\mathbf{T}$ = circle group, 8
T = nonabelian group of order 12, not A_4, not D_6, 93, 138
Torsion-free group, 175
Torsion group, 175
Torsion subgroup, 175
Transitive permutation group, 46
  doubly, 46
Translation, 5
Transposition, 31
Transvection
  matrix, 158
  transformation, 165
Transversal, 143
  Schreier, 260
Tree, 247
  maximal, 247
Turing machine, 280
Type, 202, 203

$U(n, G)$, 58
$U\{n, G\}$, 198
Unimodular matrix, 157
Unique factorization domain, 331
Unit, 61
  group of units, 130

V = Klein 4-group, 38
van Kampen, 273
Vertex set, 243

Well-defined function, 326
Well-ordered set, 329
Wielandt, 125, 135
Wilson's theorem, 10
Word, 238
  empty, 239
  freely reduced, 303
  positive, 280
  reduced, 239
  special, 286, 298
Word problem
  group, 278, 285
  semigroup, 285
Wreath product, 140

Z = integers
Z_n = integers modulo n, 20
Zassenhaus, 26
Zassenhaus lemma, 105
Zorn's lemma, 329

Rotman

THE THEORY OF GROUPS: An Introduction

Second Edition

ALLYN AND BACON

563655